Worcester Polytechnic Institute
Digital WPI

Masters Theses (All Theses, All Years) | Electronic Theses and Dissertations

2009-08-28

Analysis of the Kekkon Family in Neuronal Development

Edith Vanina Machado Plada
Worcester Polytechnic Institute

Follow this and additional works at: https://digitalcommons.wpi.edu/etd-theses

Repository Citation
Plada, Edith Vanina Machado, "*Analysis of the Kekkon Family in Neuronal Development*" (2009). *Masters Theses (All Theses, All Years)*. 995.
https://digitalcommons.wpi.edu/etd-theses/995

This thesis is brought to you for free and open access by Digital WPI. It has been accepted for inclusion in Masters Theses (All Theses, All Years) by an authorized administrator of Digital WPI. For more information, please contact wpi-etd@wpi.edu.

Analysis of the Kekkon Family in Neuronal development

by

Edith Plada

A Thesis

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Master of Science

in

Biology & Biotechnology

August 2009

APPROVED by:

Dr. Joseph Duffy	Dr. Reeta Prusty Rao	Dr. Elizabeth Ryder
Major Advisor	Committee member	Committee member
_____	_____	_____

I dedicate this work to the designer and creator of life.

May the knowledge and understanding never diminish the awe.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

A lot of people have helped and supported me in the process of this work, so I must offer my thanks and appreciation to some of them.

To Joseph Duffy, my professor and advisor, my deepest gratitude. Thank you for all the encouragement to try new things, the patience and the trust in my work and my person. I appreciate all your guidance, all the lessons taught, and all the support provided both in this project and otherwise.

To Liz Ryder. Thanks for showing me how interesting and exciting is genetic research in the nervous system. You taught me the basics of what became my chosen areas: genetics and neurobiology. I can't help thinking your teaching may have something to do with it. In fact, if it wasn't for you, I wouldn't have considered WPI. To you I'm particularly grateful.

To Reeta Prusty Rao, thank you for jumping in my committee, providing great suggestions and guidance.

To JoAnn Whitesmith, thank you for you support, encouragement and technical assistance on this project. It was a pleasure to be your TA – a very challenging and demanding one, but nevertheless a pleasure.

To Sam Politz, my appreciation for the guidance and questions answered in the process of selecting WPI for my master studies. Also thanks for the opportunity to rotate in your lab, and teaching me the basics of *C. elegans* work.

To Ally Hunter, Mike Buckholt and Jills Rulfs, thank you for the TA opportunity, instruction and guidance.

To Carol Butler, Eileen Dagostino and the administrative staff of the Biology and biotechnology department, thanks for making this possible, without you this department would crumble.

To the remaining faculty and staff, thanks for enriching this experience.

To Tzumin Lee, thanks for the feedback, opinions and heads up on this project. Your input significantly shaped this work.

To Alex Koon and James Ashley from Vivian Budnik lab at Umass, thanks for all the instruction, the protocols and tips of how to dissect 3rd instar larva.

To all my lab mates in Duffy's lab, thank you for making these last 2 years so enjoyable. Thanks for listening to me when I needed to vent, and for chatting when my brain was overheating. I already miss our lunch hours. Special thanks to Prachi Gupta, thanks for your friendship and suggesting I should seek out Duff for a rotation. I would have never joined his lab if it weren't for you.

To Gleifer Coghetto, my best friend, my lover, my husband. I love you for all the support you gave me. If you were less understanding or less encouraging I would have crumbled a long time ago. Unfortunately I can't list all the things I'm thankful for; it would make the acknowledgements longer than the thesis itself.

And to my parents, for believing in me so much, for encouraging my dreams and for understanding my busyness.

ABSTRACT

Adhesion Molecules have been associated with a number of neurological and psychological disorders (humans), and implicated in various developmental processes (animals). Better understanding the development of the nervous system and the roles of adhesion molecules in it may be crucial to better understanding these disorders. LIGs, Leucine Rich Repeat and ImmunoGlobulin containing transmembrane proteins, represent a novel class of such adhesion molecules and have been implicated in various neuronal processes, including neurite outgrowth, axonal pathfinding, neuronal regeneration and survival. Two such LIGs are Kek1 and Kek2, members of a Drosophila LIG family, which have been reported to function in axonal pathfinding and synaptic plasticity, respectively. It is unclear what their roles in these processes are, as well as if other members of the Drosophila LIG family have similar roles. Current studies aim to survey the Kekkon family function in the nervous system, looking to identify new phenotypes and/or to elucidate the mechanisms underlying previously identified phenotypes.

To achieve this goal, tissue specific inducible RNAi technique was employed. Validating of a number of transgenic RNAi stocks obtained was necessary and showed that all stocks obtained promoted specific and efficient knock down of target gene. Next an assessment of RNAi knockdown efficacy in developing nervous system was carried out and knockdown was

shown to be weak if not in the presence of Dicer-2 co-misexpression. A number of screens for general behavioral phenotypes were performed including ubiquitous, neural, and imaginal discs knockdown. These uncovered possible effects of *kek1* neural knockdown, as well as possible interaction of Kek1 with neurotactin, neuroglian and kek2. NMJ analysis of Kek5 and Kek6 was also carried out and preliminary results indicate possible interaction of kek5 in NMJ, although no statistical significance was detected.

INTRODUCTION

The nervous system is arguably the most complex biological system. Proper function relies greatly on a pattern of highly stereotyped neural projections and very precise connections formed during the development, as well as the proper transmission and receipt of signals that direct neuronal activity. As a result, enormous emphasis has been placed in understanding the molecular mechanisms that control, regulate and modulate these processes. Consequently, significant progress in elucidating these processes has been made in recent decades. Despite this progress, however, our overall understanding of neural development and brain function is still in its infancy.

NEURAL CELL ADHESION MOLECULES

At the molecular level, our understanding of the processes involved in neural development has been enhanced by the identification of neural adhesion molecules (Van Vactor, 1998). Such molecules are transmembrane proteins, often with defined extracellular motifs, that have been demonstrated to be essential for various aspects of neural development and linked to a variety of neural diseases (Katidou et al., 2008). For example, one such molecule, Neural Cell Adhesion Molecule (NCAM), is a key adhesion molecule in the vertebrate nervous system and contains 5 Immunoglobulin-like (IG) domains and 2 fibronectin (FN) type III domains within its extracellular domain (Cunningham et al., 1987). NCAM mutant mice have

diminished overall brain and olfactory bulb size, but otherwise normal nervous system structure (Campos-Ortega, 1997; Van Vactor, 1998). NCAM has been associated with various neurological disorders, such as schizophrenia, bipolar disorder, depression and anxiety disorders (Katidou et al., 2008). While no direct link has been found between genetic mutations in NCAM and the aforementioned diseases, differential regulation of NCAM isoforms has been clearly associated with these disorders. It is possible, therefore, that differences in NCAM regulation reflects a feature of these disorders, rather than represents an underlying cause. Additional links come from the use of a NCAM derived peptide for therapeutic treatment of Alzheimer's Disease (AD) due to its neuro-protective role in the pathology of AD (Klementiev et al., 2007).

The L1 family, with six IG domains, followed by three to five FN II domains, represent another important class of neural adhesion molecules (Maness and Schachner, 2007). Mutations in L1, which is known to interact with NCAM in the nervous system, has been clearly linked to an X-linked neurological syndrome of broad spectrum called CRASH syndrome (acronym for corpus collosum hypoplasia, retardation, adducted thumbs, spastic paraplegia and hydrocephalus)(Fransen et al., 1996). L1 activity may be associated with other neurological disorders as well. Fetal alcohol disorder, for instance, seems to be in part due to alcohol inhibition of some L1 functions during fetal development(Bearer, 2001). Furthermore, modification

of L1 activity has shown promising results in treatment of spinal cord injury in mouse models. Additional members of the L1 subfamily have also been associated with neural disorders, including schizophrenia, mental retardation, autism, multiple sclerosis and vulnerability to drug addiction.

LIGS REPRESENT A NEW CLASS OF NEURAL MOLECULES

In addition to the aforementioned IG and FN motifs, adhesion molecules may contain additional sequence elements governing their function. These include Leucine rich repeats (LRRs), which represent a protein-protein interaction motif consisting of a β-strand and a α-helix connected by loops, that together often form a curved, horseshoe-shaped structure. As with other NCAMs, LRR containing molecules have also been implicated in neurological disorders. For instance, SLITRK1 seems to be associated with Tourette's syndrome (Abelson et al., 2005), LGI1 is connected to temporal lobe epilepsy (Kalachikov et al., 2002) and NYCTALOPIN with congenital stationary night blindness (Bech-Hansen et al., 2000; Pusch et al., 2000).

A more recently identified class of adhesion molecules contains both **L**RRs and **IG** domains and is referred to as the **LIG** superfamily (MacLaren et al., 2004). Although individually, both the IG domain and the LRR motifs are very common, relatively few molecules contain both together and these molecules are often associated with enriched or exclusive expression in the

nervous system. Since LIGs have only recently been a focus of research, information on the functional significance of these molecules is limited (Fig. 1).

One of the most characterized sub-families of LIGs is LINGO, a family of 4 molecules with 12 LRRs and 1 IG domain (Chen et al., 2006). LINGO-1 interacts with the Nogo receptor and $p75^{NRT}$ and acts through RhoA to inhibit neurite outgrowth, and therefore has become an important target for axonal regeneration research. AMIGO/ALIVIN, another subfamily of LIGs, contains 6 LRRs and 1 IG domain and likewise has been linked to neurite outgrowth. However, AMIGO-1 (ALIVIN-2) acts to promote neurite outgrowth in vitro, in contrast to the inhibitory activity associated with LINGO-1. Also linked to neural development, AMIGO-2 (ALIVIN-1) was shown to mediate synaptic activity dependent-cell survival (Ono et al., 2003), potentially playing a role in apoptotic cell selection during neural development. AMIGO-2 also has been mapped to the

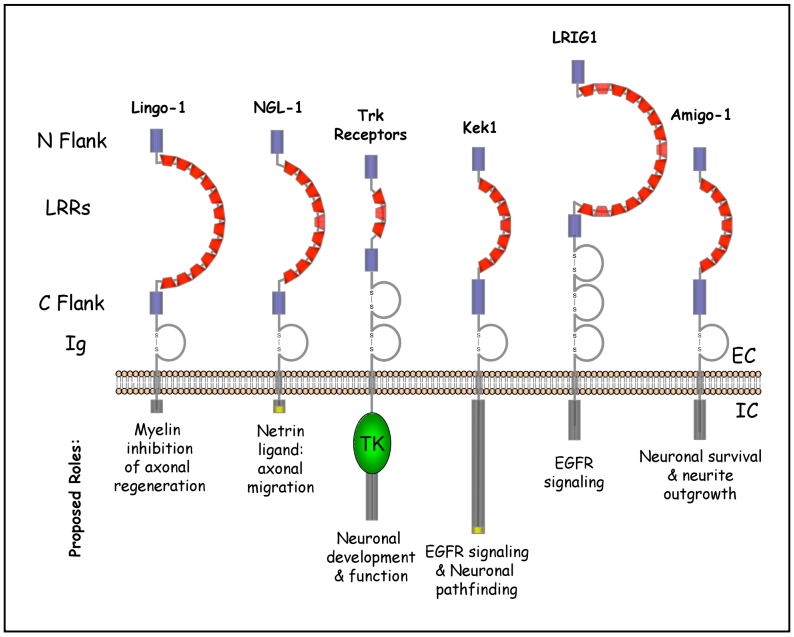


Figure 1: Members of the LIG superfamily. LIGs have variable numbers of LRRs (illustrated in red) and Ig domains (gray horseshoe), a transmembrane and an intracellular region. Only the Trks, which are receptor tyrosine kinases, have an identified enzymatic domain (green).

same region as loci associated with Alzheimer's disease type 5 and Parkinson's disease type 8 were mapped, however the molecular nature of these disease variants has not yet been determined.

Other subfamilies of LIGs include Netrin G1 ligand (NGL-1), NLRR and FLRT, all of which again appear to have associations with neural development. NGL-1 promotes growth of neurons in the thalamus during embryogenesis (Lin et al., 2003). For the NLRR family, NLRR-3 was shown to be upregulated when damage is afflicted to the brain cortex of mouse and NLRR-4 seems to have a role in hippocampal-dependent memory retention(Bando et al., 2005; Ishii et al., 1996). Finally, FLRT3 has been shown to promote neurite outgrowth and is upregulated during peripheral nerve injury(Tsuji et al., 2004).

While various links have been established between adhesion molecules and diseases, the exact mechanisms by which these molecules contribute to nervous system development and to the pathology of neural disorders is largely unknown. An improved understanding of how adhesion molecules function in the formation of the nervous system and the mechanisms through which they accomplish their various roles is an important step in unraveling the basis of neurological disorders and developing effective therapeutic strategies.

Ethical and moral complications make animal model organisms vital to the study of neuronal developmental process. The inherent tractability and

the genetic tools available in *Drosophila melanogaster*, in addition to a simplified, but largely homologous developmental mechanism, makes *Drosophila* instrumental in understanding the basic interaction of adhesion mechanisms in neural processes.

EMBRYOGENESIS

Early embryonic development of the fly is marked by the formation of a syncytium, a single multinucleated cell. Subsequently, as the embryo develops, membranes form creating a cellular blastoderm stage, and gastrulation starts along with formation of the germ band. Soon after the start of gastrulation, the germ band elongates, folding internally and extending anteriorally, thereby causing the posterior extremity of the embryo to approach the anterior extremity. Later in embryogenesis the germ band retracts until the posterior extremity of embryo reaches the posterior end

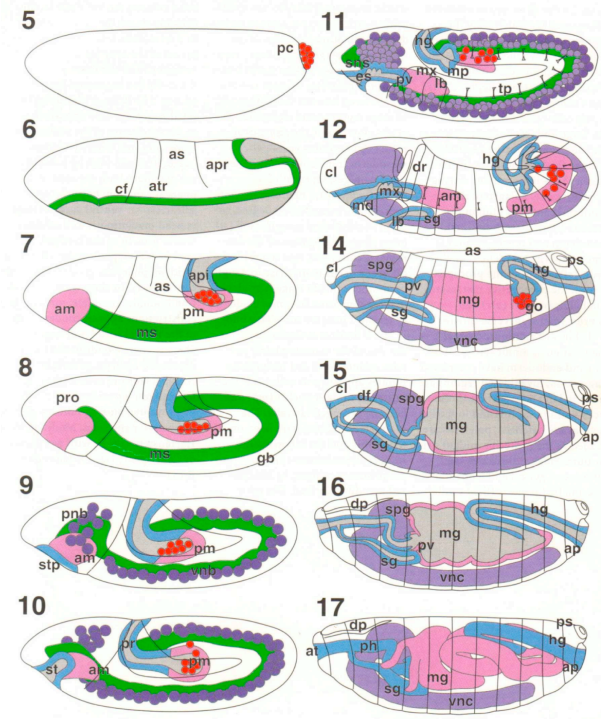


Fig 2: schematic illustration of embryonic staging and main events of embryonic development. Number at top left indicates staging. Stage 6 starts gastrulation. Stage 9 indicates neural progenitor segregation (purple) (vnb-ventral neuroblasts, pnb – procephalic neuroblasts). Stage 10 shows further segregation of neuroblasts and first neuroblast division and appearance of ganglion mother cells. Germ band elongation stops. Stage 11 epidermal segmentation becomes evident and neuroblasts division continues. Stage 12 shows ventral nerve cord entirely separated from epidermis and appearance of first neural processes and fibers. Stage 13 signifies well-differentiated ventral nerve cord and supraoesophageal ganglion and head involution begins. Fiber connectives and commissures linking the different neuromeres and muscle cells are visible at this stage. In stage 16 synapse formation starts and lasts well into time of hatching. During stage 17 ventral cord further retracts. (Campos-Ortega, 1997)

proper.

To better understand embryonic development, a classification system involving stages based on prominent features of the embryo has been developed. Figure 2 illustrates the main stages of embryonic development and its staging classification (Campos-Ortega, 1997).

NEUROGENESIS

There are three main processes in *Drosophila* neurogenesis:

1. Acquisition of neural identity,
2. Axonal guidance/fasciculation, and
3. Synapse formation.

The first process is the specification of neuroblasts followed by the formation and differentiation of neurons. Once cells have adopted a neuronal identity neurite outgrowth is initiated, thus starting the process of axonal guidance and pathfinding. The first differentiated neurons are called pioneer neurons and their axons lay out a scaffold for further development of the nervous system. Subsequently, differentiated neurons then extend their axons which fasciculate to pioneer neurons' axons. The axons of these neurons reach their appropriate target by constantly selecting and fasciculating with correct axons among many choices, a process often referred to as selective fasciculation (Goodman and Doe, 1993). The final process is the formation of synapse, which consists of three distinct steps.

The first step is appropriate synaptic target recognition, which is then followed by structural, molecular and physiological changes that characterize synapse formation. Synapses reach functional maturity by the time embryos hatch, however synapse maintenance, growth and plasticity, the third step of synapse formation, continues.

1. Acquisition of neural identity

The *Drosophila* nervous system develops in a bilateral symmetrical segmented pattern forming a sequence of repeated units called neuromeres (Campos-Ortega, 1993; Campos-Ortega, 1997). At each neuromere, clusters of neuroectodermal cells are formed and generate only one neural progenitor cell per cluster – the neuroblast. Neural fate determination is conferred by expression of proneural genes, such as members of the *achete-scute* complex, and is controlled by Delta/Notch signaling (Campos-Ortega, 1995; Duffy and Gergen, 1994). Proneural genes encode transcription factors and promote accumulation of Delta. Delta interacts with the receptor Notch on neighboring cells and induces down-regulation of the proneural genes. As a result the cell with highest levels of proneural genes expression becomes a neuroblast and delaminates from the cluster. Segmentation genes, such as the pair-rule genes *fushi tarazu* and *even skipped* and the segment polarity genes *wingless*, *hedgehog*, *patched*, *gooseberry* and *engrailed*, along with apicobasal polarity genes are also involved with specifying neuroblast

identity(Siller and Doe, 2009). Specification of a neuroblast's particular identity by the patterning genes determines its lineage, ultimately determining the identify of daughter neurons.

After a neuroblast delaminates from the ectodermal layer, it undergo several rounds of asymmetric divisions, generating a ganglion mother cell (GMC) at each division while still preserving its stem cell properties. GMCs then undergo one division to generate two distinct neurons and/or glial cells. Within *Drosophila* the number and identity of neurons generated by each neuroblast is invariant and highly reproducible.

2. Axonal guidance and fasciculation

During the second phase of neural development, the first set of differentiated neurons elongate axonal projections that are directed by guidance cues throughout the nervous system (Goodman and Doe, 1993). These cues may be released (by glial cells) and diffusible as is the case of Slits and Netrins, or they may be membrane bound molecules such as Ephrins. Semaphorins can be membrane bound or released. Guidance molecules can act as attractants or repellents depending on cell type, context and timing. Slits and Semaphorins act principally as repellent signals, and Netrin acts primarily as an attractive cue, while Ephrin can act either way.

A paradigm for axonal guidance is midline axonal crossing in *Drosophila* (Chilton, 2006; Sanchez-Soriano et al., 2007). In *Drosophila*, Slit and Netrin are both expressed and released by the midline glial cells. Expression of Slit in the midline repels growth cones thereby preventing ipsilateral projecting neurons from crossing the midline. This repulsion is mediated by Slit's receptor Roundabout (Robo), by interaction of Slit's LRR with Robo's Ig domain. In neurons that project contralaterally across the midline, however, the repulsive effect of Slit is negated by Commissureless (Comm) expression, which prevents Robo transport to the growth cone. Then the growth cone is able to cross the midline in response to Netrin attraction though the Frazzled receptor. Once growth cones cross the midline, Comm is down-regulated and Slit/Robo mediated repulsion takes place, allowing axons to proceed past the midline (Fig. 3). Independently Netrin also mediates repulsion from midline though receptor UNC-5, which is upregulated after midline crossing.

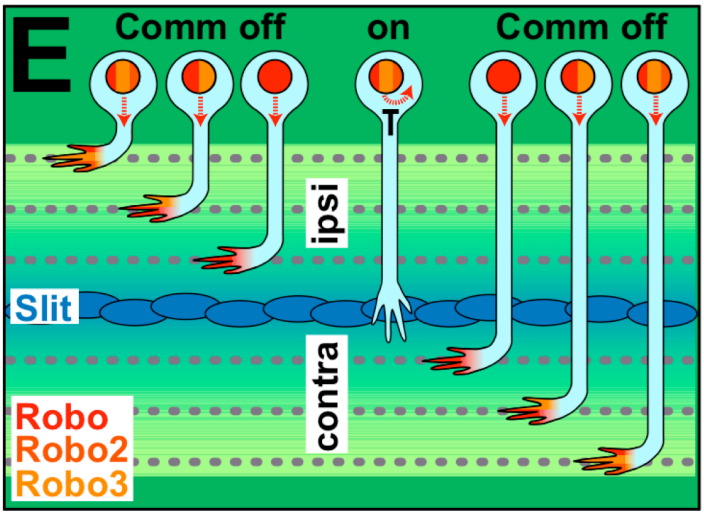


Figure 3: Illustration of Robo and Slit signaling in midline crossing. Robo expression profile determines the axonal track that axons select as they join the ventral nerve chord. Comm expression blocks Slit mediated repulsion and enable axons to cross midline. Once across midline, Comm down-regulation reinstates Slit-mediated repulsion and axon proceed to the axonal track of choice via Robo signaling. (Sanchez-Soriano et al., 2007)

Once pioneer neurons have established their axonal tracks, remaining neurons extend their axons through selective fasciculation (Goodman and Doe, 1993). Although some molecules involved in this process have been identified, the signaling mechanisms involved are not well understood. For instance FasciclinII (FasII), the *Drosophila* homolog of Neural Cell Adhesion Molecule (NCAM), has been shown to mediate axonal fasciculation. In animals overexpressing Fas II, over-fasciculation is observed to the point where axons do not defasciculate and diverge from pathways when they normally would. Furthermore, in *fasII* mutants axonal fascicles are not bundled together as tightly as wild type, indicating lack of proper fasciculation. However, no defective projections are observed in these mutants (Lin et al., 1994; Lin and Goodman, 1994). These results indicate that the mechanisms that govern selective fasciculation may be distinct from guidance cues mechanisms (Goodman and Doe, 1993; Van Vactor, 1998).

Neuroglian (NRG), a homolog of the vertebrate L1 adhesion molecule, (Bieber et al., 1989; Hall and Bieber, 1997), N-type Cadherin (DN-cadherin) (Iwai et al., 1997), matrix metalloproteinase (Miller et al., 2008) and Neurotactin (NRT) (Speicher et al., 1998) have also been implicated in axonal fasciculation. In addition, substrate adhesion molecules (SAMs), such as integrins, were shown to be involved in neuronal migration and axonal fasciculation in *C. elegans* (Baum and Garriga, 1997). Considering the more limited family of integrins in *C. elegans*, this result is believed to be

indicative of a novel role for integrins in other species as well (Baun & Garriga, 1997).

Two other major families of molecules also implicated in axonal fasciculation are receptor tyrosine kinases (RTKs), such as Derailed (Drl), and receptor protein tyrosine phosphatases (RPTPs), such as DPTP69D and DPTP99A (Van Vactor, 1998). Both kinases and phosphatases seem to modulate both axonal fasciculation as well as axonal guidance, indicating that although these processes may have separate mechanisms, they appear to share a common pathway. Figure 4 depicts several molecules involved in axonal fasciculation.

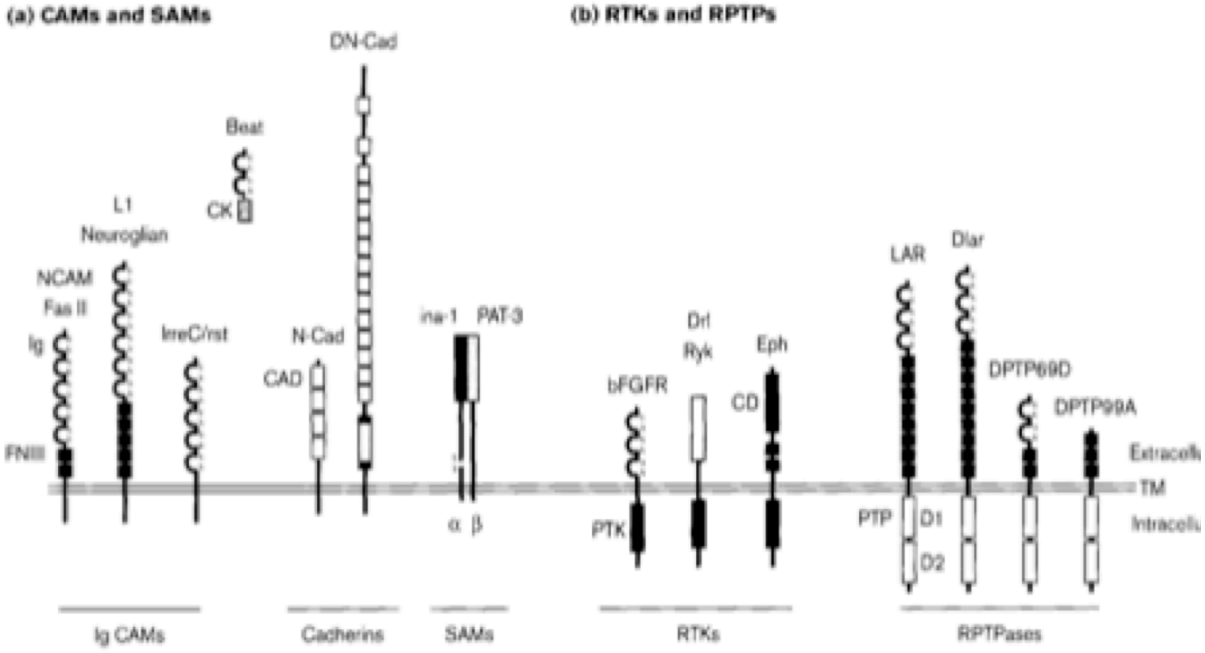


Figure 4: Molecules involved in selective fasciculation. Cell adhesion molecules featuring IG and fibronectin domains represent the most characterized molecules involved in axonal fasciculation. Substrate adhesion molecules represent a more recent class of molecules that also seems to be involved in this process. RTKs and RPTPs are seems to modulate both fasciculation and guidance. (Van Vactor, 1998)

Semaphorin Ia and Beaten Path Ia (Beat) are also believed to act as repellents in axonal fasciculation (Van Vactor, 1998). For instance, Beat Ia accumulates in high concentration on specific choice points where fascicles divide. It has been proposed that Beat Ia decreases adhesion of FasII and other CAMs to allow branching of nerves.

As a result of the mechanisms in place for axonal guidance and selective fasciculation, very clear axonal patterns are formed. Different aspects of this pattern can be observed with different markers. For instance upon staining with the antibody BP102, which targets an epitope on CNS axons, a ladder like structure can be observed in the ventral nerve cord, where the two commissures per segment can be distinguished (Fig.5A). Alternatively detecting expression of FasII in embryos with the 1D4 antibody yields a view of the three axonal tracks on each side of the midline, as well as some of the peripheral nervous system (Fig. 5B).

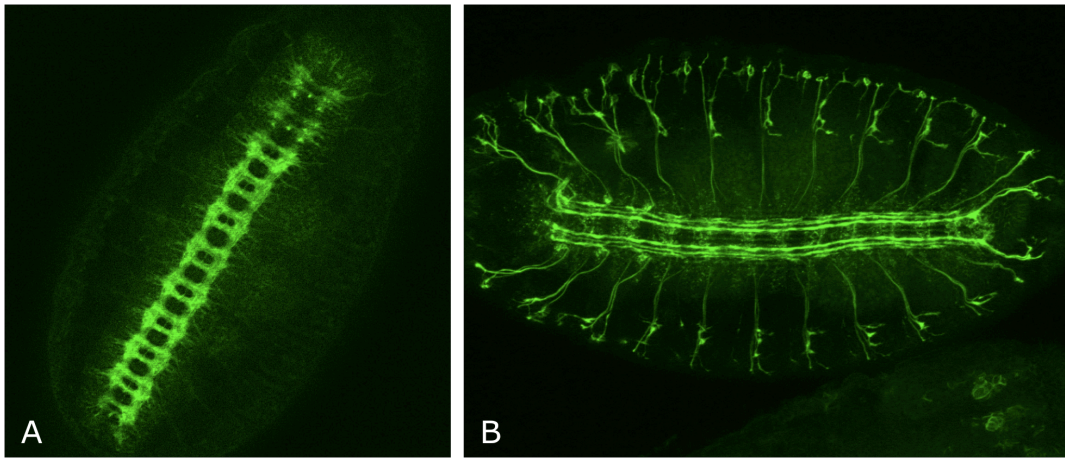


Figure 5: Axonal patterns of *Dosophila* embryo. Panel A shows CNS upon staining with BP102 antibody, which marks all CNS axons. Panel B shows staining with ID04 antibody against FASII, which marks the three major axonal tracks on each side of the midline as well as some of the peripheral nervous system. Both images were taken at stage 15/16 at 100x magnification, using Zeiss apotome processing. Right panel is a z-stack maximum image projection.

3. Synapse formation – NMJ as a model

Structure of the NMJ

The last phase of neural development involves synapse formation, the study of which is limited to a few *in vivo* models, due to the scale and complexity of synapses. Of these, the Neuromuscular Junction (NMJ) in *Drosophila* has long been used as a model for synapse formation, due to the relative ease of its manipulation and observation, as compared to synapses in the CNS. Furthermore, general aspects of NMJ formation are stereotypical and reproducible. In each hemisegment (the lateral half of each segment),

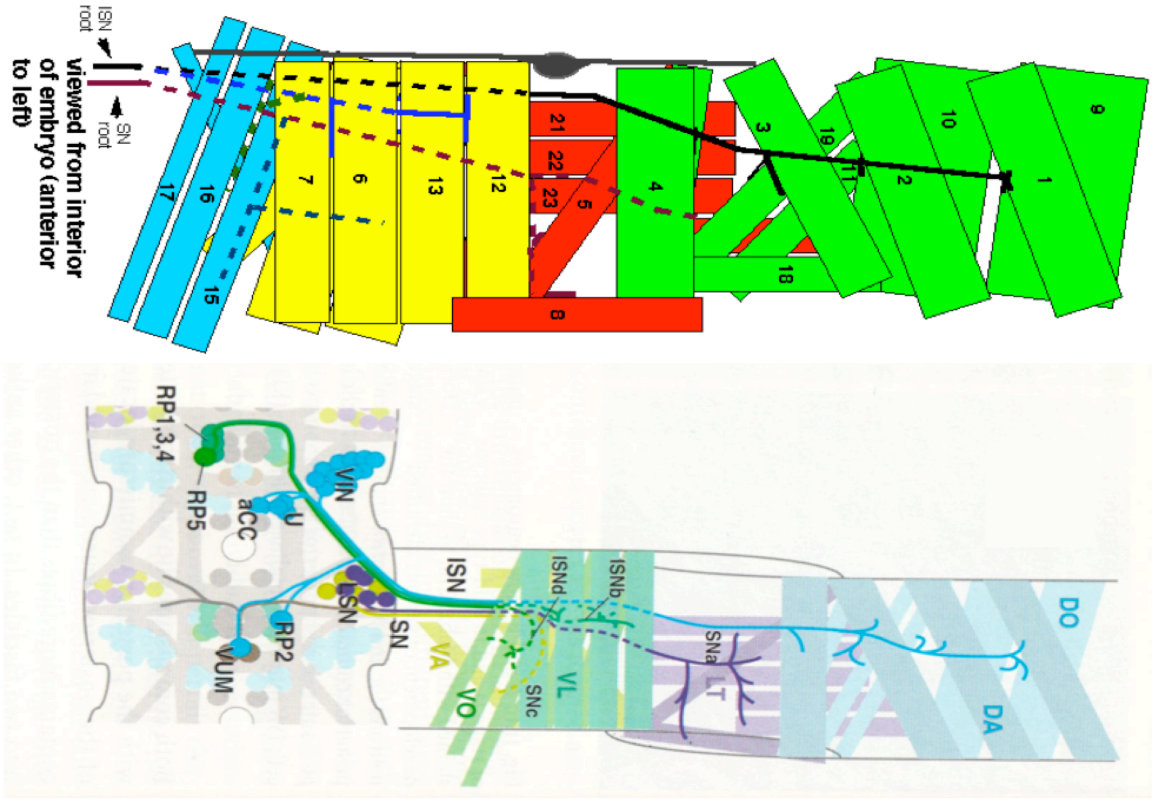


Figure 6: Muscle and motor neuron patterns. Schematic illustrations of a hemisegment muscle system. Muscle groups and nerve branches are color coded. Top panel indicates muscle pattern and muscle numbering nomenclature. Bottom panel shows section of vnc and MN cell bodies as well as labelling of nerve branches. (Campos-Ortega, 1997)

30 muscles are innervated by 33-40 motor neurons. Figure 6 shows a schematic of the observed muscle pattern; the top panel indicates the numerical nomenclature used, while the bottom panel shows the innervating neurons and their trajectories through the various nerve branches.

Generally speaking, each motor neuron (MN) innervates one or more muscle fiber and forms a specific type of synapse, Ib, Is, II, or III. Since only a few neurons have been identified to innervate specific muscle fibers, synapses are identified by bouton size. Each fiber is innervated by only one neuron of each type (Hoang and Chiba, 2001). Structurally, type Ib (for big) boutons are the largest, measuring 3-6μm; they are glutamatergic, are present in all muscles and tend to be in short and minimally branched terminals. The majority of MNs form type Ib boutons and innervate only one or immediately adjacent muscle fibers (as is the case of muscle 6/7, 21-24 and 15-17). Type Is (small) boutons are slightly smaller than Ib, about 2-4μm, also glutamatergic and exist in longer and more elaborate terminals. They are believed to be present in all muscles, though in some cases the exact bouton type present is not clear. Type II boutons are small (1 to 2μm), they use glutamate and octopamine as neurotransmitters, and are present in most muscle in very long and elaborate terminals. The innervation pattern of Is and II are similar, with one MN innervating multiple muscle fibers in a given muscle group. Finally, type III boutons are of a medium size, about 2-

3μm, are believed to be present only on muscle 12, and contain both glutamate and insulin.

TARGET RECOGNITION AND SYNAPTOGENESIS

A key aspect of synapse formation involves target recognition. Axonal guidance mechanisms collaborate in achieving synaptic specificity as they lead the growth cone toward the appropriate synaptic target. The final step before synapse formation is the selection of the target cell from among its many neighbors. Filopodial processes in presynaptic growth cones probe the environment to identify the appropriate target, following the same strategy as used during axonal pathfinding. Likewise, filopodia-like processes are also present in postsynaptic cells, both in dendrites and muscles (Ritzenthaler and Chiba, 2001). These processes in muscles are called myopodia and they enable direct and dynamic interaction between the growth cone and possible muscle targets at relatively long distances. The interaction between filopodia and myopodia is sufficient for target recognition and formation of a stable connection between growth cone and muscle.

Target identification is believed to occur through molecules with very specific expression patterns. Arguably, one of the most well-characterized examples is FasIII (Rose and Chiba, 2000). During synapse formation within the PNS, FasIII is expressed in the RP3 motor neurons and its synaptic targets - muscles 6 and 7. In *fasIII* null mutants, incorrect innervation is observed, albeit at low frequency – 9% of RP3 neurons form connections

with other muscles. In Fas3 overexpression in muscles, however, incorrect innervation is observed in 72% of the cases, indicating that gain-of-function experiments can have stronger phenotypes than loss-of-function experiments, most likely because of redundant mechanisms that exist to ensure precise connectivity. When FasIII is overexpressed in muscles in a null background (no FasIII in RP3 neurons), the frequency of incorrect innervation decreases to 14%, demonstrating that phenotype is potentially conferred primarily through a homophilic interaction.

Another important molecule linked to synapse formation is the LRR containing immune receptor Toll, which is also a synaptic repellent molecule expressed in muscles 15, 16, 17 and 29, a muscle group that the RP3 growth cone crosses on its way to M6/7 (Rose and Chiba, 1999; Rose et al., 1997). In a *toll* knockout, only 10% of RP3 growth cones reach the appropriate target. Upon misexpression of Toll in muscle, only 14% of RP3 growth cones form appropriate connections, while the other 86% form incorrect innervations or no innervations at all. Interestingly, misexpression of both Toll and Fas3, repellent and attractant respectively, in muscle results in few defects, indicating that growth cones receive and integrate multiple inputs to generate a decision regarding synaptic target recognition.

Likewise, temporal regulation is also key, as expression of Toll in proximal muscles ensures that inappropriate synapses are not formed by MNs that will innervate more distal muscles. Subsequently, Toll is then

down-regulated by the time the appropriate MNs should synapse in these proximal muscles.

Not surprisingly, ligands (Wnt4, Netrin), receptors (Frizzled and Derailed), LRR containing adhesion molecules (Capricious) and transcription factors have all been implicated in specific aspects of synaptogenesis. From work on these and other molecules, it has become clear that synaptic specificity is not only conferred by attractant signals, but also by repellent signals from surrounding inappropriate targets. In support of this latter notion, several synaptogenic inhibitory molecules have been identified to date, such as Dishevelled, Beaten path, and D-semaphorin. Moreover, it appears that some molecules may have dual roles. For instance, Netrin and Wnt-4 have been shown to have synaptogenic and anti-synaptogenic properties in a context dependent manner. Such effects are likely mediated by distinct receptors as observed in axonal guidance.

One model proposes that synaptogenesis occurs through two different classes of molecules (Hoang and Chiba, 1999). The first class includes molecules with very narrow expression patterns and promotes specific target recognition, such as FasIII. Manipulation of these molecules would generate very specific single cell level abnormalities. The second class includes molecules with broader expression profiles that promote synaptogenesis after initial target recognition steps. Manipulation of these molecules would cause more generalized defects, although these could be subtle due to

possible functional redundancy of molecules of this class. Consistent with this model, there are several molecules identified as general adhesion molecules that promote synaptic formation in a general manner, including Neuroglian, Integrin and Dn-Cadherin.

Upon target recognition, the relatively flat growth cone swells and forms large prevaricosities, which then constrict to form smaller varicosities or boutons, typical of a mature synaptic connection. At this point in time, clusters of presynaptic and postsynaptic apparatus is formed and co-localized across the synaptic cleft in distinct active zones. It is largely unknown how adhesion molecules promote the molecular changes associated with synaptogenesis, and what are the components/pathways involved in this process. It is known however that pre- and postsynaptic growth is tightly regulated and coordinated, particularly during larval development, when muscles grow to become about 150X their original volume.

A predictable series of events has been observed during the development of the synapse between muscle 6 and the RP3 motor neuron. Prior to synaptogenesis, neither neurotransmitter nor a functional receptor is present on MN or muscle. As the motor neuron growth cone contacts the muscle, transmitter expression starts, and myotubes uncouple soon after. Immediately after uncoupling of myotubes, a small number of functional glutamate receptors are evenly distributed on the muscle surface. Shortly

after motor neuron filopodia localize at the developing synaptic zone, functional receptor localization occurs. At this point a functional synapse forms, endogenous muscle activity begins and nerve stimulation leads to muscle contraction. Then presynaptic specialization develops, giving rise to the mature morphology. After that a second motor neuron contacts muscle 6 at the pre-established synaptic zone, a second stage of functional receptor expression emerges and vigorous neuromuscular activity characteristic of larval locomotory movement initiates.

Presynaptic activity, although not required for initial synapse formation is required for post synaptic clustering of Glutamate Receptors (GlutR) and regulates structure and strength during NMJ growth (Budnik, 1996; Nakayama et al., 2006; Prokop and Meinertzhagen, 2006). Pre and postsynaptic FasII levels are also tightly coregulated (Ashley et al., 2005). In fact, upregulating FasII activity only on one side of the synapse decreases synaptic size and bouton number and generates abnormal NMJ morphology. However, equal upregulation of FasII on both pre and post-synaptic cells increases NMJ size significantly, stimulating new bouton formation.

An additional mechanism to coordinate pre and postsynaptic growth is retrograde Bone Morphogenic Protein (BMP) signaling (Keshishian and Kim, 2004; Nakayama et al., 2006; Prokop and Meinertzhagen, 2006). BMP signaling is a conserved signaling cascade that controls many developmental processes. In *Drosophila,* members of the BMP pathway are required for

proper NMJ development and manipulations that decrease BMP signaling negatively affects synapse size, stability and homeostasis. It is believed that type II receptor Wishful thinking (Wit) as well as both type I receptors, Saxphone (Sax) and Thickvein (Tkv) are involved in retrograde signaling at the NMJ (Keshishian and Kim, 2004). Specifically, the ligand Glass bottom boat (Gbb) is released by the postsynaptic muscle terminal onto the synaptic cleft, which promotes phosphorylaton of the downstream transcription factor Mad in the motor presynaptic terminal (Goold and Davis, 2007). BMP component mutants, such as BMP receptor *wishful thinking* (*wit*), show decreased NMJ size and bouton number to approximately 40% of wild type (Aberle et al., 2002). On the other hand, overexpression of BMP components or mutants of negative regulators of BMP components exhibit larger NMJ size. For instance *highwire* (*hiw*) mutant, a putative E3 ligase that negatively regulates levels of co-Smad Medea shows synaptic bouton number increase of up to 200% over wild type (Aberle et al., 2002; McCabe et al., 2004).

THE KEKKON FAMILY

Our lab has been studying a set of proteins called the Kekkon family. This class of LIG family members are transmembrane proteins identified in invertebrates with seven leucine-rich repeats and an Immunoglobulin domain (Fig.7) (MacLaren et al., 2004). Embryonic expression profiles for family members indicates they are expressed in the nervous system,

however, to date little is known about the functional role of these molecules in neural development.

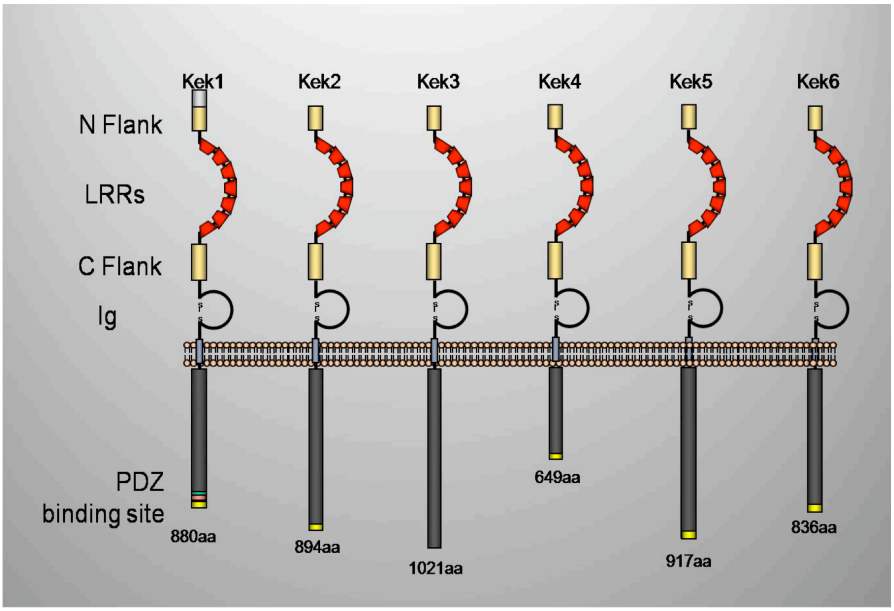


Figure 7: The Kekkon Family. Constituted of 6 transmembrane molecules containing 7 LRR and one Ig domain. Their intracellular region does not contain a catalytic activity domain but most display a PDZ binding site.

Kekkon1 (Kek1) was initially identified on the basis of its expression in the nervous system (Musacchio and Perrimon, 1996) and has subsequently been shown to interact with Epidermal Growth Factor receptor (EGFR) as a negative regulator (Ghiglione et al., 1999). In addition, in a study of *neurotactin (nrt)* role in axonal fasciculation, the *nrt* and *kek1* double mutant shows ventral nerve cord fasciculation defects not seen in the single mutants, indicating a possible role of *kek1* in axonal fasciculation (Speicher et al., 1998). Synergistic genetic interaction were also observed between *nrt* and *neuroglian* and *derailed,* but not with *DPTP69D* and *DPTP99A*. However, no mechanisms for these interactions have been proposed, leaving open the significance of these findings.

Kekkon5, another member of the Kek family, has recently been demonstrated to modulate Bone Morphogenesis protein (BMP) signaling in

the crossvein development in the *Drosophila* wing (Evans et al., 2009). Misexpression and loss-of-function of *kek5* was shown to affect the profile of phosphorylated Mad and dSRF in presumptive crossvein cells. Furthermore, Kek5 phenotypes are similar to those obtained by manipulation of Short gastrulation (Sog), a secreted modulator of BMP signaling, but unlike phenotypes of dominant negative receptors, indicating Kek5 may be acting upstream of BMP receptors. Additionally Kek5 was shown to antagonize Glass bottom boat (gbb), a BMP ligand, supporting this claim.

Recently, Kekkon2 was identified in a microarray screen for genes involved in synaptic plasticity (Guan et al., 2005). Kek2 was upregulated and downregulated in mutants in which synaptic activity levels were increased and decreased long-term, respectively. It was then shown that the absolute level of *kek2* expression modulates the extent of innervation in the NMJ, where both increase and decrease in Kek2 levels caused a decrease of bouton number in the NMJ (30-50% decrease), thereby supporting a role for *kek2* in synaptic plasticity.

The goal of my thesis research project was to perform a broad survey of the role of the Kek family in neural development. For family members this included expression profiling, functional tests using RNAi-mediated knockdown, and attempts to reproduce the reported phenotypes of axonal fasciculation and synaptic plasticity to further characterize these interactions and possibly determine specificity of the involved Kek family member.

RESULTS

KNOCKDOWN STRATEGY

To enable a survey of the function of Kek family members in *Drosophila*, RNA interference (RNAi) was used. In *Drosophila*, transgenic hairpin constructs capable of producing gene specific RNAi triggers can be coupled to the existing GAL4/UAS system to promote inducible knockdown of genes in specific tissues (Duffy, 2002). Manny lines have already been established which induce expression of the GAL4 transcription factor in a variety of tissues and patterns. When expressed, GAL4 binds its recognition site – upstream activating sequences (UAS), promoting expression of the desired target sequence. This system is often used to promote misexpression of genes in a tissue of interest, by inserting the coding region of the genes of interest downstream of the UAS recognition site. In this bipartite system, UAS responder lines are created, and then flies are mated with GAL4 drivers of interest to promote expression of the gene of interest.

Recently, the Vienna Drosophila RNAi Center (VDRC) has created a library of transgenic stocks with inducible RNAi constructs targeting 88.2% of the *Drosophila* genome (Fig.8) (Dietzl et al., 2007). In this library, the UAS promoter is attached to an inverted repeat sequence of approximately 300-400bp that matches the target gene. Upon transcription, the inverted repeat folds over creating a hairpin RNA (hpRNA). The hpRNA is recognized by the RNA interference machinery of the cell and the Dicer enzyme (Dcr-2 in *Drosophila*) cuts the hpRNA into short stretches of RNA of about 19-23bps each, thereby forming a set of silencing RNAs (siRNA) all matching the target mRNA (Ghildiyal and Zamore, 2009). With the help of Dcr-2 and the double stranded RNA binding protein R2D2 the siRNAs are subsequently loaded onto the RNA-induced silencing complex (RISC), which with the help of Argonaute2 (Ago2) selects the guide strand from the siRNA.

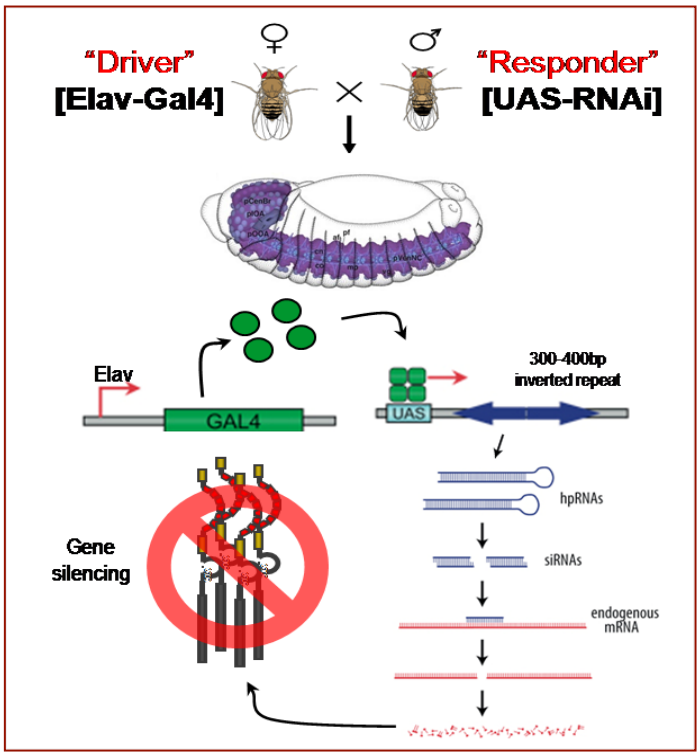


Figure 8: The GAL4-UAS.RNAi system. Parental flies have driver constructs directing GAL4 expression in tissue specific manner (e.g. nervous system), or a responder construct that directs RNAi trigger transcription under control of UAS sites. Progeny express GAL4 in the nervous system, which then directs transcription of the RNAi trigger in that tissue. RNAi trigger forms hpRNAi, which is processed by Dicer-2, generating siRNAs. The gene specific siRNAs then couple with the RISC complex to promote degradation of the target gene mRNA, thereby silencing target gene expression.

At this stage, Ago2 cleaves the passenger strand of the siRNA, forming mature RISC loaded with a single stranded RNA (guide strand). This guide strand then directs identification of mRNAs that have complementary sequence, promoting their degradation, and effectively preventing translation of target mRNA.

In theory, then, with the VDRC strains and existing GAL4 lines, the function of ~88% of the genes in the *Drosophila* genome can quickly be assessed in the tissue of interest (Fig. 8). Transgenic RNAi responder lines targeting each of the *keks* were available and obtained from VDRC. The sequences used to generate the hairpin for each *kek* family member are shown in Appendix 2. To examine further the published genetic interactions between Kek1 and Neurotactin (Nrt) and to study the possibility of an interaction between Neuroglian (Nrg) and Kek1, VDRC RNAi lines targeting *nrt* and *nrg* were also obtained.

A significant concern when using RNAi mediated knockdown is possible OFF target effects; that is adverse effects caused by knockdown of non-target genes whose sequence matches possible siRNA sequences derived from the hairpin sequence, but which are not the intended target. To address this concern, the VDRC has created a scoring system to evaluate their constructs for possible off-target effects (Dietzl et al., 2007). For each line created, the number of ON targets is published, as well as the number of OFF targets, and a relative specificity ranking called the S19 score.

Definition of an ON target is any gene that contains a perfect match to at least 50% of the 19-mer siRNAs generated by a hairpin construct. In contrast, an OFF target is defined as any gene that contains sequence similarity with at least one 19-mer, but less than 50% of all 19-mers possibly generated by the construct. To calculate a the S19 relative specificity score, the following formula is used:

S19 = Σ ON target matches /(Σ ON target matches + Σ OFF target matches).

Hence, in a line that has only one ON target and no OFF target, S19 equals 1. This is the most desirable scenario to insure that only the desired target gene is knocked down, and it is the case with most of the VDRC lines obtained for this work. One exception is *kek5* RNAi strain 47770, which has one ON target and 5 OFF target genes. However, the calculated S19 for this line is 0.98, which means that although there are 5 OFF targets, 98% of the possible 19-mer siRNAs produced are complementary to the *kek5* mRNA, while only 2% of the 19-mer siRNAs have complementarity with OFF targets. Moreover, for the five OFF target genes, four genes are targeted by only a single 19-mer siRNA and the fifth gene is only targeted by three possible 19-mer siRNAs. This is in contrast to *kek5*, which is targeted by all of the possible 290 19-mer siRNAs produced by the hairpin. This indicates that even though OFF target effects are possible for this line, such effects are unlikely to reduce expression of the OFF target genes to a level that would

be physiologically significant. Table 1 lists all lines obtained from VDRC, including number of OFF targets and S19 score for each line.

Table 1: RNAi lines obtained from VDRC and validation data

Target Gene	Line #	Inserted chromo.	OFF target	S19	PCR	GFP	Phen.	Tested against
Kek 1	36252	3	0	1	✓	✓	✓	-
Kek 1	43521	2	0	1	✓	✓	✓	Kek 6
Kek 1	4761	2	0	1	-	✓	✓	-
Kek 2	42449	2	1	1	-	✓	-	Kek 1, 6
Kek 3	6354	3	0	1	✓	-	-	-
Kek 3	6356	2	0	1	✓	-	-	-
Kek 4	915	2	0	1	-	✓	-	-
Kek 5	27249	1	5	0.98	✓	✓	✓	-
Kek 5	47770	2	5	0.98	✓	✓	✓	Kek 1, 2, 4
Kek 6	27164	0	0	1	✓	✓	-	Kek 1
Kek 6	27165	0	0	1	✓	✓	-	Kek 1, 2, 5
NRT	8495	2	0	1	✓	-	-	-
NRG	27201	3	0	1	✓	-	-	-

OFF target - indicates number of OFF target genes for each line
S19 - score calculated as Σ ON matches / (Σ ON matches + Σ OFF matches)
PCR - indicates result of construct validation by PCR
GFP - result of functional validation by knockdown of GFP-tagged target gene
Phen. - functional validation by suppression of misexpression phenotype of target gene
✓ - indicates positive validation was obtained
"**-**" - indicates no validation data was obtained
Tested against - indicates target genes against which indicated RNAi line was tested as a negative controls in functional validation assays. No cross activity was observed with any controls tested

RNAi Line Validation

Since the project was largely reliant on the RNAi lines obtained, it was crucial to confirm their identity and efficacy. Initially, to verify the presence of the indicated construct, PCR was performed using a primer designed to

match a short sequence in the pUAST transgene vector within the UAS binding region and a gene specific primer to amplify a section of the construct excluding the inverted repeat. Lines that were validated by use of PCR are also indicated in Table 1. To verify that the lines were capable of effecting silencing of the desired target gene, functional validation was also carried out. This was assayed in two ways. First, the RNAi lines were crossed to strains in which a GFP tagged version of the desired target gene is expressed in the eye using the GMRGAL4 driver. The presence of the RNAi trigger should then lead to degradation of the mRNA for the GFP tagged target gene resulting in a loss of GFP fluorescence in the adult eye. For *keks 1,2,4,5,* and *6*, gene specific loss of GFP fluorescence was observed, thereby demonstrating that each respective RNAi line is capable of effectively reducing expression of the desired target gene (Fig. 9 and 10 and Table 1). In addition, RNAi effects are limited to single family members with no cross family effects observed (Fig. 9, Table 1). Because of the lack of a GFP-tagged version this assay was not carried out for *kek3*.

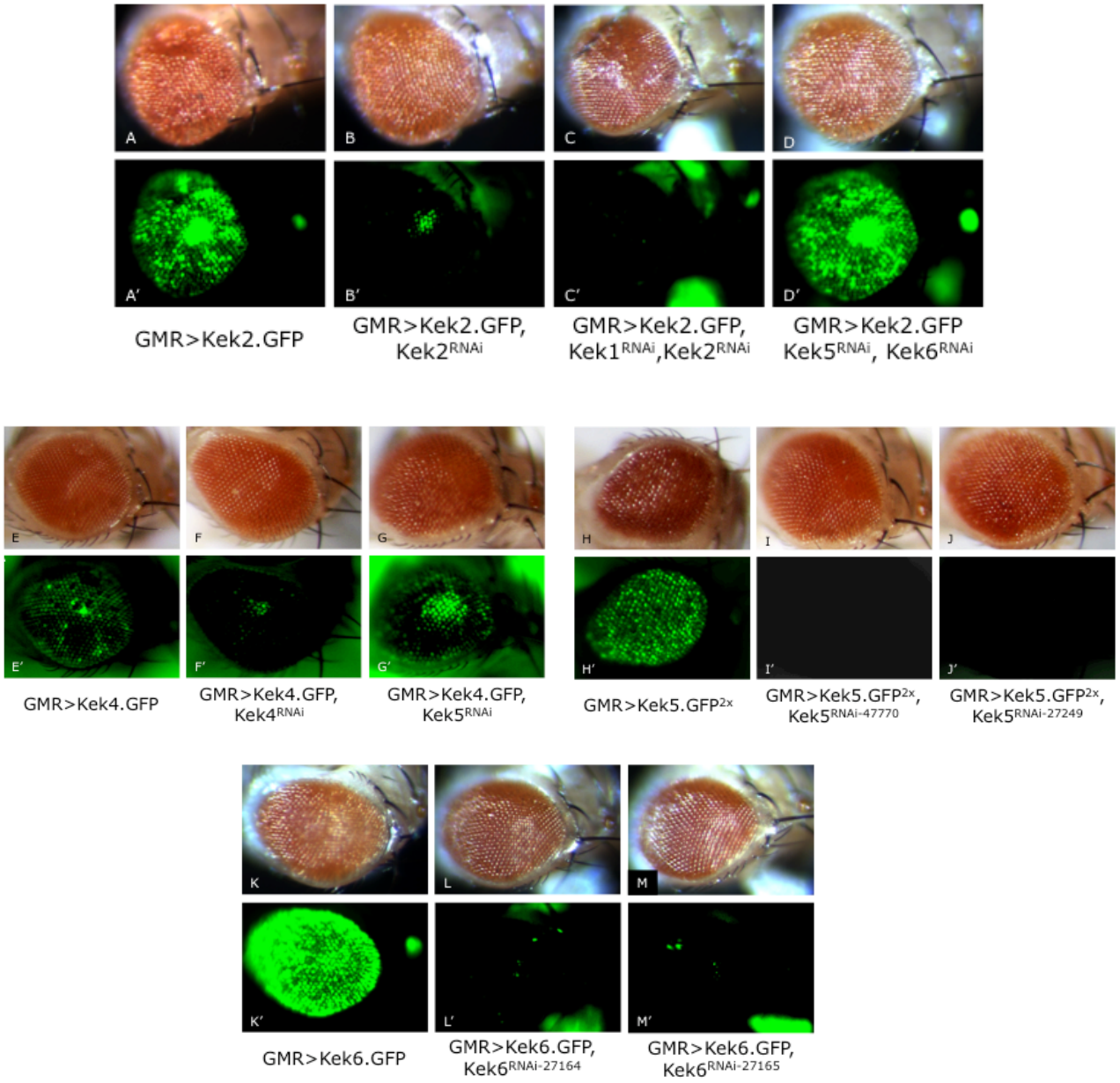


Figure 9: Functional validation of *kek* family RNAi lines through Kek-GFP knockdown. Brightfield (A-M) and epifluorescent (A'-M') micrographs of adult compound eyes. Upon GMRGAL4 mediated misexpression of Kek-GFPs, significant GFP expression is observed in the adult eye (panels A', E', H' and K'. Introducing the presence of gene specific RNAi results in efficient knockdown of appropriate target gene as observed by the loss of Kek-GFP expression in the adult eye (B', C', F', I',J', K' and L').

Another method to functionally validate the lines obtained was through suppression of misexpression phenotypes. In the case of Kek1 and Kek5, misexpression has known phenotypes. Therefore, RNAi lines could be validated by their ability to suppress the misexpression phenotypes when co-expressed with the target gene (Table 1). Misexpression of Kek1 in the eye with the GMRGAL4 driver causes a rough eye phenotype and *kek1* RNAi lines were able to fully suppress this phenotype (Fig. 10). Likewise, *kek5* RNAi lines suppress the effects of misexpression of Kek5 in the wing (severe wing blisters) by the apterousGAL4 (apGAL4) driver (Fig. 10).

In the case of Kek1, less GFP is detected in the eye, relative to the other Keks, even in the absence of the RNAi trigger. This is due to inhibition of the EGFR by misexpression of GFP-tagged Kek1, which leads to a loss of the Kek1-GFP expressing photoreceptor cells and thus the rough eye phenotype (Alvarado et al., 2004).

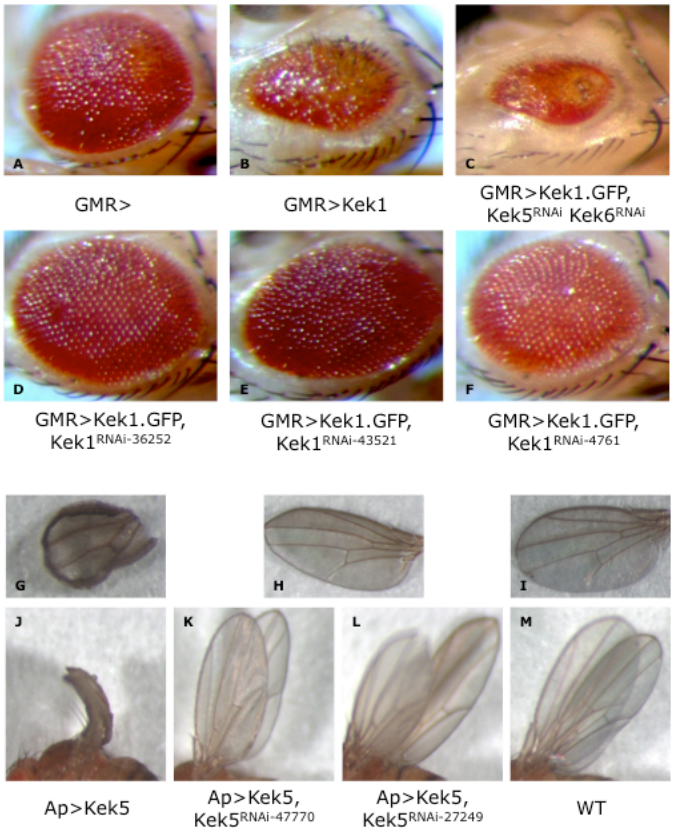


Figure 10: Functional validation of RNAi lines by suppression of misexpression phenotypes. Brightfield micrographs of adult compound eyes (A-F) and wings (G-M). GAL4 mediated misexpression of Kek1 and Kek5 leads to rough eye and abnormal wing phenotypes, respectively (B, G, J). Introducing the presence of gene specific RNAi results in suppression of these phenotypes (D-F, H, K, and L).

However, coexpression of the *kek1* RNAi hairpin triggers knockdown of Kek1-GFP expression, thereby restoring EGFR activity and the presence of photoreceptor cells as observed by suppression of rough eye phenotype. However, although photoreceptor cells are restored, GFP expression is not present also confirming knockdown of the *kek1-GFP* mRNA (data not shown). Thus, all of the assayed *kek* family RNAi lines appear to promote efficient, gene specific knockdown in the tissues tested. It should be noted however, some non-specific effects are observed, as expression of RNAi lines using apGAL4 driver appears to promote a held out wing phenotype.

DEVELOPMENTAL ASSESSMENT OF RNAI KNOCKDOWN EFFICACY

The validation tests above confirmed effectiveness and specificity of the RNAi lines in the developing eye and wing. However, the effectiveness of transgenic RNAi in the *Drosophila* nervous system has not been addressed to date. Moreover, communication with others in the field indicated that the efficacy of transgenic RNAi in the embryonic nervous system was questionable.

Hence, I aimed to validate the use of the transgenic RNAi technique in this context. To that end we expressed GFP tagged Kek5 in the nervous system using the pan neural C155GAL4 driver, with and without co-expression of the *kek5* RNAi trigger. Levels of GFP in the nervous system in embryo, 1st instar and 3rd instar larva were then compared among the

genotypes. Some knockdown was observed in the presence of RNAi; knockdown was more robust in 3rd instar larva than in embryos or 1st instar, but not as significant as might be necessary in order to carry out functional studies (Fig. 11). Therefore, I attempted to improve efficiency of knockdown either by expressing 2 copies of *kek5* RNAi trigger and/or by co-expressing Dcr2. Expressing 2 copies of Kek5 RNAi gave a minimal increase in knockdown efficiency. One possibility is that Dcr2, which is required to generate siRNA triggers from the hairpin RNA, is limiting in the nervous system. If so, then increasing Dcr2 levels should lead to increased knockdown effects in the nervous system. Consistent with this, significant knockdown of Kek5-GFP was observed in the presence of Dcr2 co-expression (Fig. 11). It was observed that the co-expression of Dcr2 increases the variability of Kek5-GFP expression levels in the nervous system. Furthermore, even in the presence of Dcr2 and RNAi, GFP expression varied and very often some GFP can still be observed in the embryo and 1st instar larva, indicating that knockdown of Kek5-GFP is not complete in early stages of development. In contrast, GFP levels in dissected 3rd instar larval brains were not variable and seemed to indicate complete GFP knockdown.

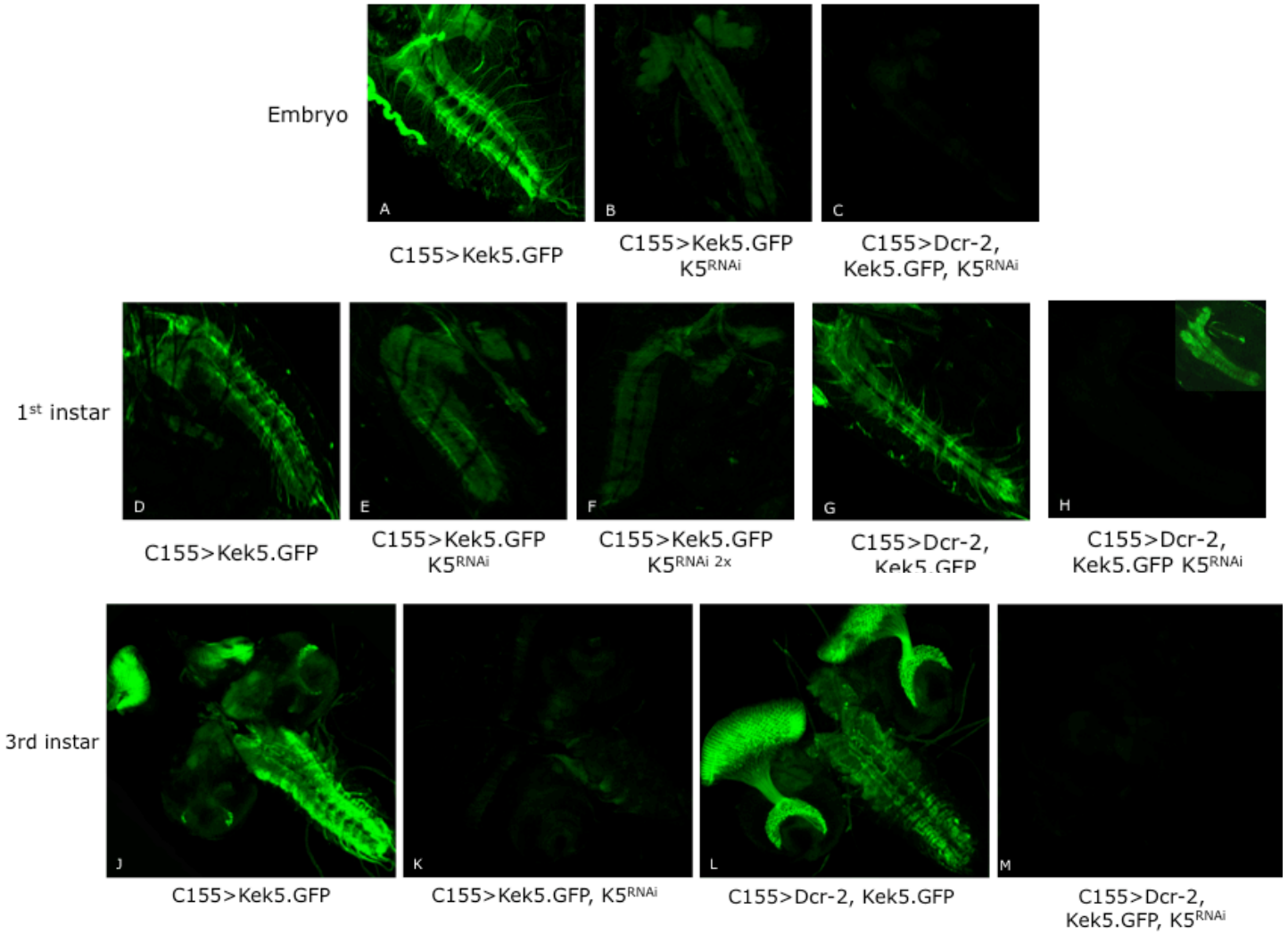


Figure 11: Assessment of RNAi knockdown in developing nervous system. Epifluorescent micrographs of the nervous system of embryos (A-C), early 1st instar larva (D-I) and 3rd instar larva (J-M). Bright GFP can be observed in the ventral nerve cord when Kek5-GFP is misexpressed by C155GAL4 (A, D, J). Some, albeit not robust, knockdown is observed when Kek5 RNAi triggers are also expressed (B, E, K). Having two RNAi trigger constructs is not sufficient to increase knockdown significantly (F). Adding Dicer does slightly decrease GFP intensity in the absence of RNAi (G, L) but in the presence of RNAi, it enhances knockdown appreciably (C, H, M). However some GFP is still visible in early development even in the presence of Dcr and RNAi. All images are taken with similar exposure and adjusted to similar brightness, contrast and color levels, except inset in H, which has been adjusted in brightness and contrast to become visible.

CREATING COMBINATORIAL KNOCKDOWN STRAINS

Overt neuronal phenotypes have not yet been reported for single mutants of Kek1 and Kek5. In addition, Kek2 phenotypes reported in NMJ are quite subtle. It is believed that functional redundancy in the nervous system is prevalent which may be particularly relevant in the study of the Kek family function since the Keks share strong structural homology. Hence, combinatorial knockdown may be an important component of the Kek functional investigation. Prior phylogenetic analyses in the lab revealed that the Kek family is divided into 2 clades where Keks1, 2 and 3 form a clade separate from Keks4, 5 and 6 (MacLaren et al., 2004) (Fig. 12), suggesting an initial combination for multiple knockdowns. Furthermore, this phylogenetic analysis shows that although the Kek family has been evolutionarily conserved for approximately 500 million years, Kek4 arose

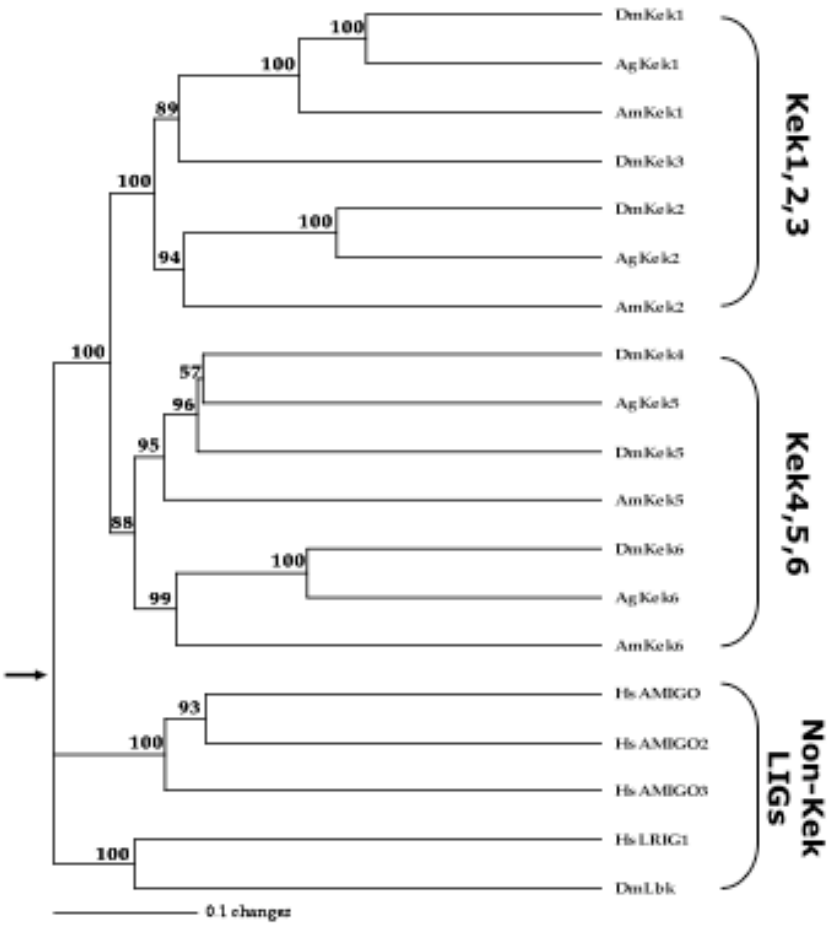


Figure 12. Phylogenetic analysis of Kek family (from T. Evans)

more recently than that, and Kek3 has been lost in some species after its emergence (*A. mellifera* and *A. gambiae*; Fig. 13), indicating they may be more dispensable for the organism.

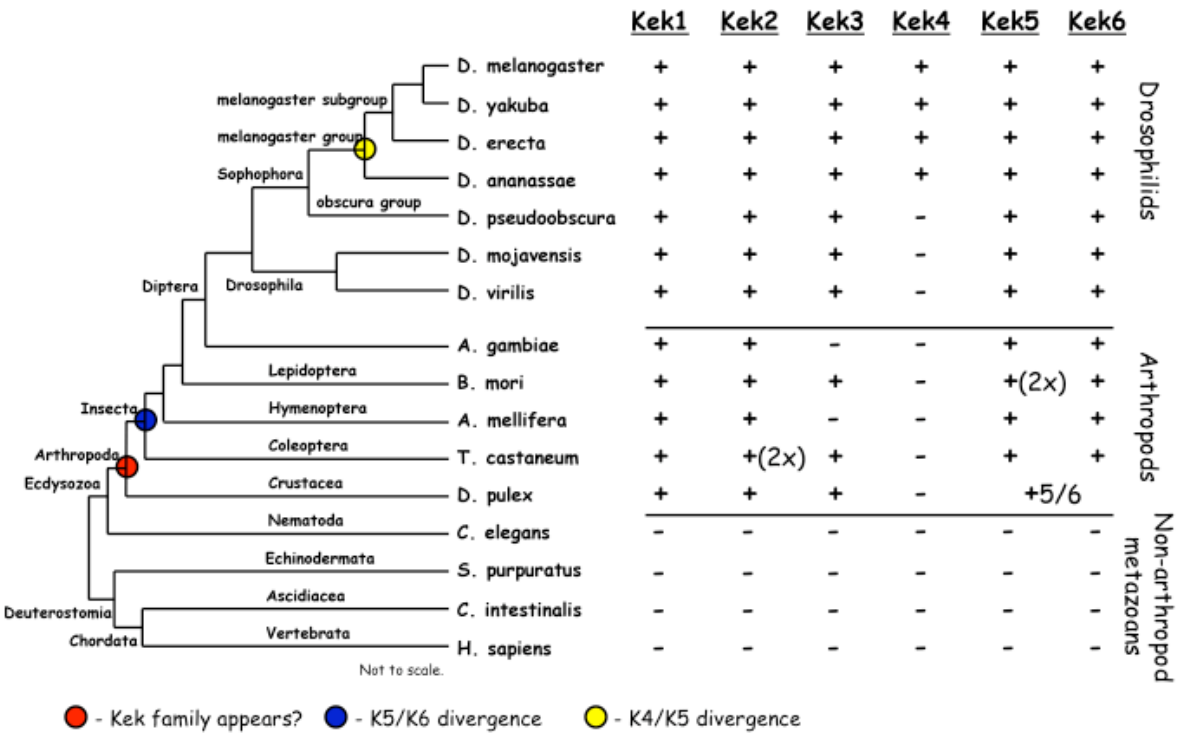


Figure 13. Analysis of Kek family evolution (from T. Evans).

Therefore, as an initial strategy, lines were created to generate pairwise knockdowns for the most well conserved loci in the two clades, namely a *kek1* and *kek2* recombinant and a *kek5 and kek6* recombinant. These lines were validated using similar strategies as the original RNAi lines as is summarized in Table 2. Recombinant lines were also created with *kek1, nrt, nrg* and *kek2* to further investigate the reported interaction between *kek1* and *nrt* and to study the specificity of this interaction.

Table 2: Recombinant RNAi lines created and validation data

Target Genes	Line #	1st Line	2nd Line	Chrom.	PCR		GFP		Phen.	
					1st	2nd	1st	2nd	1st	2nd
Kek1 Kek2	43	43521	42449	2	✓	-	-	✓	✓	-
Kek1 Kek2	47	4761	42449	2	-	-	-	✓	✓	-
Kek5 Kek6	64	47770	27164	2	✓	✓	✓	✓	✓	-
Kek5 Kek6	65	47770	27165	2	✓	✓	✓	✓	✓	-
Kek1 NRT	43	43521	8495	2	✓	✓	✓	-	✓	-
Kek1 NRT	47	4761	8495	2	-	✓	✓	-	✓	-
Kek1 NRG		36252	27201	3	✓	✓	✓	-	✓	-
Kek2 NRT		42449	8495	2	-	✓	✓	-	-	-

PCR - indicates result of construct validation by PCR
GFP - result of functional validation by knockdown of GFP-tagged target gene
Phen. - functional validation by suppression of misexpression phenotype of target gene
✓ - indicates positive validation was obtained
"**-**" - indicates no validation data was obtained

PRIMARY SCREEN FOR KEK FAMILY FUNCTION

Ubiquitous knockdown

Expression of the *kek* family has been demonstrated to be principally in neural tissue, but expression in additional tissues has also been observed. For example, *kek1* expression has been reported in the wing and ovary, in addition to the eye, where it is a downstream target of the EGFR, while *kek5* expression appears to be more ubiquitious with some enrichment in the embryonic nervous system. Therefore, to assess the effect of global knockdown, the role of each *kek* in the organism was assessed using two ubiquitous drivers, tubulinGAL4 (tubGAL4) and actinGAL4 (actGAL4). No

overt phenotype was observed other than reduced viability, which was observed for Kek1, Kek3 and Kek5 (Tables 3 and 4). Viability for each cross was calculated based on the following formula:

(UAS-*kek*^{RNAi}/GAL4 driver) / #(UAS-RNAi/Balancer), n equals the total number of flies scored. Although percent viability for each *kek* family member varied between both drivers, effects with tubGAL4 appear more severe with lower viability numbers.

Table 3: Screen of Tubulin driven RNAi knockdown

Target Gene	Line #	Viability	n	Observation
Kek1	43521	0%	55	
Kek1	4761	0%	52	
Kek2	42449	215%	41	
Kek3	6354	0%	45	pupal lethality
Kek3	6356	0%	58	pupal lethality
Kek4	915	159%	75	
Kek5	27249	5%	22	
Kek5	47770	4%	29	
Kek6	27164	120%	8	
Kek6	27165	104%	47	
Kek5 Kek6	64V	50%	57	
Kek5 Kek6	65V	53%	26	
Kek5 Kek6	64IV	29%	79	

Table 4: Screen of Actin driven RNAi knockdown

Target Gene	Line #	Viability	n	Observation
Kek1	43521	2%	59	Walk up deficiency
Kek1	4761	2%	86	
Kek2	42449	109%	90	
Kek4	915	118%	61	
Kek5	27249	29%	40	Walk up deficiency
Kek5	47770	15%	38	
Kek6	27164	106%	55	Jumpy
Kek6	27165	80%	88	

Based on the RNAi results, *kek1, kek3,* and *kek5* appear to be vital loci, essential for viability, while no major effects are seen for *keks 2,4* and *6*, with the latter results arguing against lethality as a non-specific result of RNAi induction. For *kek5*, the reduced viability with RNAi is consistent with reported data for a *kek5* null mutant, albeit more severe in the case of the tubGAL driver. In contrast, the lethality observed with the *kek1* RNAi was somewhat unexpected as deletion of *kek1* is viable (Musacchio and Perrimon, 1996). Although this raises suspicion over the 0% viability observed with the tubGAL4 driver, the relative viability of the *kek1* null mutant has not been reported and thus could indeed be low. Similarly, although circumstantial evidence suggests that loss of *kek3* is not lethal, *kek3* null mutants have not yet been isolated, so it remains possible that knockdown of *kek3* could lead to effects on viability. Consistent with this, lethality observed with the *kek3* RNAi occurred during the pupal stage, which was not observed in *kek1* RNAi, supporting the possibility of a distinct functions for these genes.

It is important to note that for both *kek1* and *kek3* there appear to be no OFF target genes based on the VDRC scoring system. However, if lethality associated with GAL4 mediated knockdowns is non-specific - caused by soaking up essential cellular components of the RNAi processing machinery, increasing levels of this component would restore levels, thereby yielding more accurate results. Since Dcr2 seems to be a limiting factor for

RNAi processing in the nervous system, it is possible that it may be a limiting factor in other contexts as well. Therefore, knockdowns of *kek1* and *kek3* were tested again using actGAL4 in the presence of additional Dcr2 (genotypes - *UASkek1*^{RNAi}, *actGAL4 UASdcr2* and *UASkek3*^{RNAi}, *actGAL4 UASdcr2*). In addition, an RNAi trigger for GFP, which is not present in the *Drosophila* genome and for which knockdown should have no effect, was used as a control. Surprisingly, in the presence of additional Dcr2, no RNAi expressing flies were recovered, including the GFP RNAi control (although it's not clear if OFF targets for the GFP RNAi hairpin exist). The number of control (balancer) flies recovered was 15, 21, 31 and 34 for *UASkek1*$^{RNAi-4761}$, *UASkek3*$^{RNAi-6354}$, *UASkek3*$^{RNAi-6356}$ and *UASGFP*$^{RNAi-9331}$ respectively. Although only a small number of progeny were scored, the data obtained with GFP$^{RNAi-9331}$ indicates that simply activating the RNAi mechanism in the presence of Dcr2 misexpression by actGAL4 greatly reduces viability.

Thus, although RNAi can trigger gene specific knockdown with no obvious non-specific or detrimental effects in the eye, interpretation of results with ubiquitious drivers appears much more complex.

Neuronal knockdown I

Given that the results with a tissue specific driver (GMRGal4) provided clear-cut results and expression of *kek* family members is principally in the nervous system, I limited the remainder of my RNAi studies to tissue-specific

drivers. For this I utilized pan-neural drivers C155Gal4 and ElavGAL4, which express Gal4 under the control of a gene, *elav*, expressed predominantly in all neurons. Initially, knockdown of *keks* in the nervous system was performed without the presence of additional Dcr2 as an enhancer. Flies were then assayed at the cellular and organismal levels. First, the overall pattern of the embryonic nervous system was assessed with antibodies that identify the longitudinals and commissures (anti-FasII and BP104). Using these antibodies, no obvious defects were uncovered for any of the *keks* tested, as well as for the *kek5/kek6* and *kek1/nrt* recombinants (Fig. 5 and data not shown).

At the organismal level, three assays (tap, vortex, and flight) were used to evaluate the behavior of adult flies. In the tap assay flies are tapped down in the vial and their behavior observed. Wild type flies climb up the vial immediately after being tapped down. The vortex assay was performed to evaluate flies response to physical stress. After wild type flies in a vial are vortexed for 5 seconds they quickly recover – regain balance, move up, and groom. Finally, a flight assay was done to evaluate flies' flight response. Upon being dropped into a graduated cylinder, wild type flies fly and alight onto the walls, without falling to the bottom.

After running all the adult assays, no significant behavioral anomaly was observed. It was noted that RNAi expressing flies groomed less, particularly after vortex assay, and that they jumped around more.

However, this phenotype was observed in all genotypes and does not seem to reflect a specific gene function. Nonetheless, in view of my results indicating the limited RNAi knockdown efficiency in developing nervous system in the absence of additional Dcr2, I can only conclude that the *keks* are not essential for basic adult nervous system function. This conclusion is supported in part by another study in the lab that indicates that without additional dicer, RNAi mediated knockdown may be effective in the nervous system as early as late pupae stage. In this case it can also be concluded that the *Keks* function is also not vital in late pupal development.

Viability data was also obtained using ElavGAL4 driver in a similar method as viability data was obtained for ubiquitous drivers. This line has the same expression profile as C155, but it was observed to induce lower levels of responder expression. Slightly decreased viability was observed for a few lines, particularly Kek4, with no effects observed for GFP RNAi trigger in contrast to the viability results obtained with the ubiquitious drivers (Table 5, data not shown).

Table 5: Viability data for Elav driven RNAi knockdown

Target gene	Line #	Viability	n
Kek1	36252	148%	82
Kek1	43521	79%	93
Kek1	4761	89%	70
Kek2	42449	102%	107
Kek3	6354	142%	29
Kek3	6356	111%	154
Kek4	915	66%	128
Kek5	47770	111%	74
Kek5	27249	103%	201
Kek6	27164	235%	124
Kek6	27165	81%	107

Neuronal knockdown II: addition of Dicer2

To better address the role of *keks* during neural development, RNAi lines were also crossed to a strain containing both the C155GAL4 driver and UASDcr2. Since the driver and most RNAi lines were homozygous, the lack of control classes in the progeny precluded the ability to obtain accurate viability data for most lines (Table 6). However, all crosses (including *wild type* and *UASGFP*^{RNAi} controls) were set up with the same number of parental males and females of approximately the same age. Thus, between cross comparisons provide an approximate estimate of the number of progeny that should be recovered. Although not standard, in the absence of sib control classes this comparison provides one simple measure for determining if large effects on viability occur. For lines without control sibs, estimated viability was calculated as: # gene specific RNAi progeny/116 (the average for the two control crosses). Where Balancer (control) sibs were present and standard percent viability could be calculated, the % of control recovered should be approximately half of the viability for that line.

Although the results of the behavioral assays were not quantified, behavioral observations were noted. In particular it was observed that in a few lines most larva pupated near the food or in the food itself. Upon increases in levels of the molting hormone ecdysone, wild type larva crawl out of the food and pupate up on the sides of the vial. Inability to do this may denote a larval behavioral deficit, either in locomotor activity, the

ecdysone pathway, or in environment sensing. This was most severe in the *nrg*, *kek1* and *kek1/nrg* recombinant lines, and also occurred in the *kek5/kek6* recombinant line, albeit less severely. Other observations worth noting were that in the *nrg/kek1* recombinant knockdown, flies did not walk up the vial normally and were visibly uncoordinated, which likely resulted in the large numbers of adults stuck in the food. Large numbers of adults stuck in the food, consistent with a locomotion deficit, was also observed in other lines, including the *nrg* and *kek1* individual knockdowns. Additionally, in the *kek1/kek2* combinatorial knockdown, flies did not walk up the vial normally and appeared visibly uncoordinated. Although no clear phenotype was observed for any of the single *kek* knockdowns, *kek1* knockdown shows evidence of a locomotor deficit and decreased viability. Furthermore, 2 of the *kek1* recombinant knockdowns (*kek1/nrg* and *kek1/kek2*) demonstrated increased locomotor deficit. It is still unclear if the increased deficit is due to a synergistic interaction or an additive effect. Either way, this is the first time that any interaction has been observed between *kek1* and *kek2* or *kek1* and *nrg.* Likewise, this is also the first time that any indications of a neuronal phenotype has been reported for single *kek1* knockdown. In comparison no phenotype was observed for the other *keks*. Nonetheless, this could be a result of inefficient knockdown early in development or due to a lack of specific and sensitive assays to detect behavioral anomalies.

Table 6: Screen of neuronal RNAi knockdown with Dicer missexpression

Target Gene	Line #	RNAi	Balancer	Viability	% control	% males	% Stuck	Observations
K1	35252	70			60%	39%	6%	-
K1	43521	23			20%	0%	100%	PCF
K1	4761	34			29%	26%	62%	PCF
K2	42449	89			76%	56%	2%	-
K3	6354	83			71%	40%	5%	-
K3	6356	47			40%	19%	4%	1 missing 1/2 thorax
K4	915	72			62%	52%	1%	-
K5	27249	80			69%	50%	3%	-
K5	47770	111			95%	42%	0%	-
K6 (cyo)	27164	41	53	77%	35%	44%	0%	-
K6	27165	56			48%	57%	0%	-
Lambik	42570	88			76%	50%	2%	-
LIG	7993	113			97%	44%	0%	-
NRG	27201	37			32%	46%	32%	MPCF, 2 CW, 40% SU, uncoordinated
NRT	8495	52	44	118%	45%	35%	21%	-
GFP	9331	111			95%	36%	10%	-
K1K2	47I	61			52%	6%	20%	PCF, 1 curly wing
K1K2	47II	24			21%	13%	58%	PCF
K1K2	43II	10			9%	0%	60%	MPCF, very uncoordinated
K1K2	43III	5			4%	0%	80%	MPCF, very uncoordinated
K1 NRT	47I	47			40%	15%	57%	MPCF
K1 NRT	47II	41			35%	7%	10%	PCF, 1 SU
K1 NRT (cyo)	43I	2	82	2%	2%	N/A	100%	PCF
K1NRG (TM3)	I	0	37	0%	0%	N/A	N/A	PCF, uncoordicated
K1NRG (TM3)	II	15	N/A		13%	0%	67%	MPCF, uncoordicated
K1NRG (TM3)	III	8	40	20%	7%	0%	0%	MPCF, uncoordicated
K5K6 (cyo)	64IV	32	27	119%	27%	23%	19%	PCF
K5K6	64V	63			54%	38%	16%	PCF
K5K6	65II	53			45%	25%	0%	PCF
K5K6	65IV	49			42%	27%	2%	PCF
WT	W1118	122			105%	48%	2%	-

Average control progeny # 116.5 Control groups are W1118 and GFP^{RNAi}

Legend: PCF = some pupae cases close to or in food SU = stuck unfurled

MPCF = most pupae cases in food itself CW = curly wing

Imaginal discs knockdown

One of the best-characterized members of the Kek family is Kek5, which has been shown to interact with BMP signaling in wing crossvein development. Mutants in *kek5* exhibit crossvein defects, both in anterior (ACV) and posterior crossvein (PCV), in approximately 30% of adults. Defects include ectopic crossvein around ACV, truncated, missing or meandering PCV and ectopic PCV material. Crossvein signaling is also disrupted when *kek5* is misexpressed using engrailedGAL4 (enGAL4), a segment polarity gene that drives expression on the posterior compartment of each segment, including the region of presumptive ACV and PCV in the developing wing. This allowed me to explore the efficacy of RNAi in a different tissue for which a *kek* null phenotype has been well characterized. RNAi lines against *kek5*, *kek6* and their recombinant were crossed to enGAL4. Wings were then scored for crossvein defects.

Some defects were observed with enGAL4 and the *kek5* RNAi line, including ectopic ACV material and truncated PCV (Table 7). Knockdown with one *kek6* RNAi line gave a high penetrance of crossvein defects, but not in the other. In this combinatorial knockdown, defects were mostly missing or truncated ACV. When both *kek5* and *kek6* were simultaneously knocked down, defects in ACV increased to 100%. Moreover, the percent of missing ACV increased as compared to *kek6* alone and defects in PCV also increased

when compared to *kek5* alone, supporting a possible interaction between *kek5* and *kek6*.

Table 7: Crossvein Screen using En driven RNAi knockdown

Target Gene	Line #	n	% defects	
			ACV	PCV
Kek5	47770	98	1	5
Kek6	65	112	69	0
Kek6	64	88	6	1
Kek5 Kek6	65 II	61	100	44
Kek5 Kek6	64V	148	1	3

However, additional results cause concern regarding the validity of this phenotype. To start with, the frequency of defects observed with the *kek5* line is significantly lower than that observed with the null allele (20-30% missing or truncated PCV). Second, only one of the *kek6* lines presented any phenotype or interaction with *kek5*, while no significant phenotype or interaction was observed in the other *kek6* RNAi line. This could be explained by a difference in expression levels of the corresponding *UASkek6*^{RNAi} lines due to position effects on the transgenes, however both lines were equally effective in knocking down Kek6-GFP expression in the eye with GMRGal4. *kek5*^{RNAi}, *kek6*^{RNAi} and the *kek5/kek6* RNAi recombinants were also tested with the neural driver ScabrousGAL4 and no overt phenotypes were observed.

Neuronal Misexpression

Although the true benchmark for characterizing gene function *in vivo* is the analysis of loss-of-function effects, important information can often be derived from gain-of-function studies as well. Thus, as a complement to the RNAi approach to assessing Kek function in the nervous system, I also performed a screen in which members of the *kek* family were individually misexpressed in neural tissues, with the exception of *kek3* for which no transgenic misexpression lines exist. Specifically, *kek1, 2, 4, 5* and *6* were misexpressed with the pan-neural driver C155GAL4. No overt abnormalities in adult behavior were observed upon misexpression of the single responders. However, misexpression of *kek5* with a strain containing two responders led to an extremely high percentage of flies stuck in the food (82%) and flies that were not stuck on the food showed clear behavioral defects: they did not climb up the vial, walk around, groom, attempt flight or display any movement typical of wild type flies and typically were dead in a few days. In addition, many flies had wings that remained uninflated, consistent with a lack of motor coordination since wings require stroking for inflation. Thus, Kek5 expression must be appropriately regulated for wild type neural development to occur.

To gain better insight into the mechanism underlying the misexpression effect of Kek5, an intracellular variant - $Kek5^{ICΔ123}$, which lacks three conserved motifs within the intracellular domain, was also tested.

Misexpression of this Kek5 variant appears to have increased activity in other misexpression assays, producing phenotypes often more severe than 2X misexpression of wild type Kek5. Similar effects to 2X wild type Kek5 misexpression were observed with respect to viability and flies stuck in the food, although all stuck flies had uninflated wings. In contrast, however, the $Kek5^{IC\Delta 123}$ flies that did not get stuck exhibited behavior much more typical of wild type flies than those with 2X Kek5 misexpression. Hence, it seems that at least in the context of neural misexpression this variant does not produce increased activity relative to wild type Kek5.

NMJ ANALYSIS

No clear and overt phenotype was observed in the general screen, but it is possible that *kek* neural phenotypes are highly specific and would only be detected at the tissue or cellular level. For instance a *kek2* phenotype has only been observed in the NMJ (Guan et al., 2005). *kek5* was also reported to be downregulated by chronically increased synaptic activity by the same study that reported a *kek2* function in synaptic structure. Furthermore, Kek5's interaction with BMP signaling, a known modulator of synaptic structure, makes it a likely effector in elaborating NMJ structure and would be a plausible explanation for the overt behavioral phenotypes observed with high levels of Kek5 misexpression. Thus, given the technical demands of NMJ analysis and the putative likelihood of uncovering a role for Kek5, I focused

primarily on Kek5 in the interest of uncovering a novel phenotype. Analysis was carried out in *kek5* null mutant, *kek5* RNAi-mediated neural knockdown (with Dcr2 coexpression), and misexpression of both $Kek5^{WT}$ and $Kek5^{IC\Delta123}$. To test the hypothesis of functional complementation, Kek5/Kek6 combinatorial knockdown was also tested, as well as Kek6 RNAi neural knockdown (with Dicer) and neural misexpression.

Initial analysis of the *kek5* null mutant, $kek5^{fe148}$, suggested muscle pattern and innervation defects. Defects in the uncoupling of muscles 6 and 7 were observed in 2 out of 5 larva, and complete lack of innervation of muscle 6 and 7 was also observed in the same larvae, in a total of 3 hemi-segments. However, similar defects were not observed in larva from a *kek5* null mutation with a different genetic background ($Df(1)JA27/kek5^{fe148}$) and were occasionally observed in wild type larva, indicating that this is unlikely to represent a *kek5* phenotype.

In the *Drosophila* neurobiology field, standard analysis of NMJ structure relies on quantification of boutons and is often normalized over muscle surface area. Thus, quantification of boutons in the NMJ of muscles 6 and 7 in segment A2 was carried out at 10x magnification (representative images in Fig. 14). Value from both sides were averaged and analysis of variance (ANOVA) was carried out among genotypes comparing raw values of bouton number, as well as bouton number normalized over muscle surface area (Table 8 and 9 respectively). Because of developmental

correlations between physiological activity and overall muscle size, I also tested bouton number normalized by average muscle length.

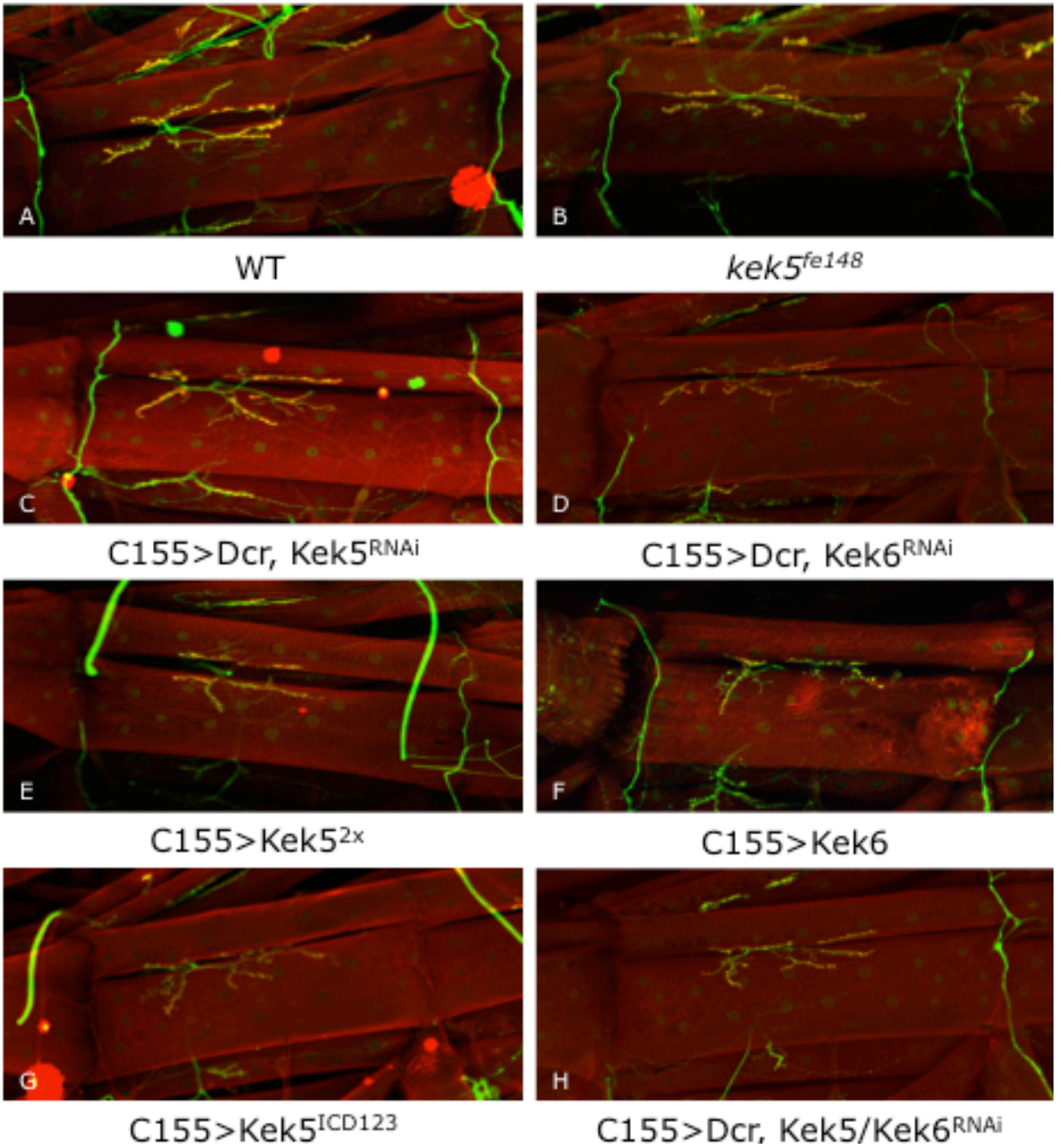


Figure 14: Representative images of NMJ. Epifluorescent micrographs of the NMJ of muscle 6/7 of 3rd instar larva for indicated genotypes. Anti-HRP (nervous system) stained in green and post-synaptic marker, Discs large, in red. Images captured at 10x with apotome processing.

Table 8: Analysis of Variance for bouton number

Summary

Groups	*Sample size*	*Sum*	*Mean*	*Mean relative to wild type*	*Variance*
W1118	17	2264.5	133	100.0%	496
K5Fe148	10	1151.5	115	86.4%	1236
C155>Dcr,K5RNAi	10	1520.5	152	114.1%	423
C155>Dcr,K5K6RNAi	8	1160.5	145	97.9%	542
C155>K52x	12	1667.495	139	102.3%	690
C155>K5IC?123	5	681.5	136	114.0%	1111
C155>Dcr,K6RNAi	6	911.5	152	106.7%	1335
C155>K6	4	568.5	142	108.9%	770

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*p-level*	*F crit*
Between Groups	9239.1	7	1319.882	1.772	10.84%	2.1564
Within Groups	47671.7	64	744.8709			
Total	56910.9	71				

Table 9: Analysis of Variance for bouton per muscle surface area

Summary

Groups	*Sample size*	*Sum*	*Mean*	*Mean relative to wild type*	*Variance*
W1118	17	0.0279	0.0016	100.0%	0.
K5Fe148	10	0.0168	0.0017	102.6%	0.
C155>Dcr,K5RNAi	10	0.0178	0.0018	108.7%	0.
C155>Dcr,K5K6RNAi	8	0.015	0.0019	97.2%	0.
C155>K52x	12	0.0206	0.0017	114.0%	0.
C155>K5IC?123	5	0.0093	0.0019	109.8%	0.
C155>Dcr,K6RNAi	6	0.0108	0.0018	92.5%	0.
C155>K6	4	0.0061	0.0015	114.7%	0.

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*p-level*	*F crit*
Between Groups	0.	7	0.	0.5106	82.32%	2.1564
Within Groups	0.	64	0.			
Total	0.	71				

Table 10: Analysis of Variance for bouton per muscle length

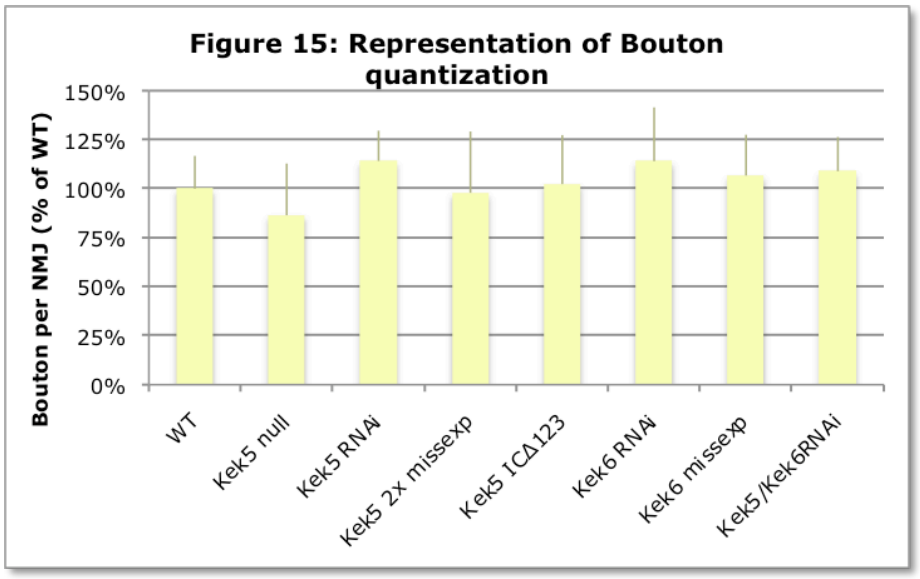
Summary

Groups	*Sample size*	*Sum*	*Mean*	*Mean relative to wild type*	*Variance*
W1118	17	4.8277	0.284	99.9%	0.0026
K5Fe148	10	2.6092	0.2609	91.9%	0.0082
C155>Dcr,K5RNAi	10	3.1399	0.314	110.6%	0.0017
C155>Dcr,K5K6RNAi	8	2.534	0.3167	97.0%	0.0065
C155>K52x	12	3.6693	0.3058	107.4%	0.0028
C155>K5IC?123	5	1.5245	0.3049	110.6%	0.0054
C155>Dcr,K6RNAi	6	1.8846	0.3141	104.0%	0.0044
C155>K6	4	1.1817	0.2954	111.5%	0.0034

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*p-level*	*F crit*
Between Groups	0.0249	7	0.0036	0.8697	53.53%	2.1564

No statistically significant difference was detected in NMJ size, irrespective of the normalization performed. Trends were observed, particularly in the number of boutons per NMJ (Fig. 15). However their relevance is unclear, for example, while bouton number in the *kek5* null is approximately 14% lower than wild type, RNAi knockdown of *kek5* resulted in an increase in the number of boutons (approximately 14%).



Although we might expect similar trends in both genotypes, it is important to note that there is a tissue-specific difference in the two loss-of-function approaches. In the null allele, *kek5* activity is removed both pre and post synaptically, while in the RNAi knockdown it was only removed pre-synaptically. Thus, it is possible that relative levels of Kek5 in neural tissue and muscle tissue are important for proper NMJ size and structure. In fact, this is exactly the case for FasII and maybe consistent with the role of BMP signaling in pre and post-synaptic co-regulation. Although no effect was observed in response to Kek5 misexpression pre-synaptically, there was a general impression that NMJ structure seemed less developed with smaller boutons. The quantification of bouton size was not carried out, but indicates that further analysis is warranted to establish Kek5 function in NMJ development. Furthermore, higher resolution analysis (such as 40x magnification) may be required to obtain more accurate and possibly less variable results (Fig. 16).

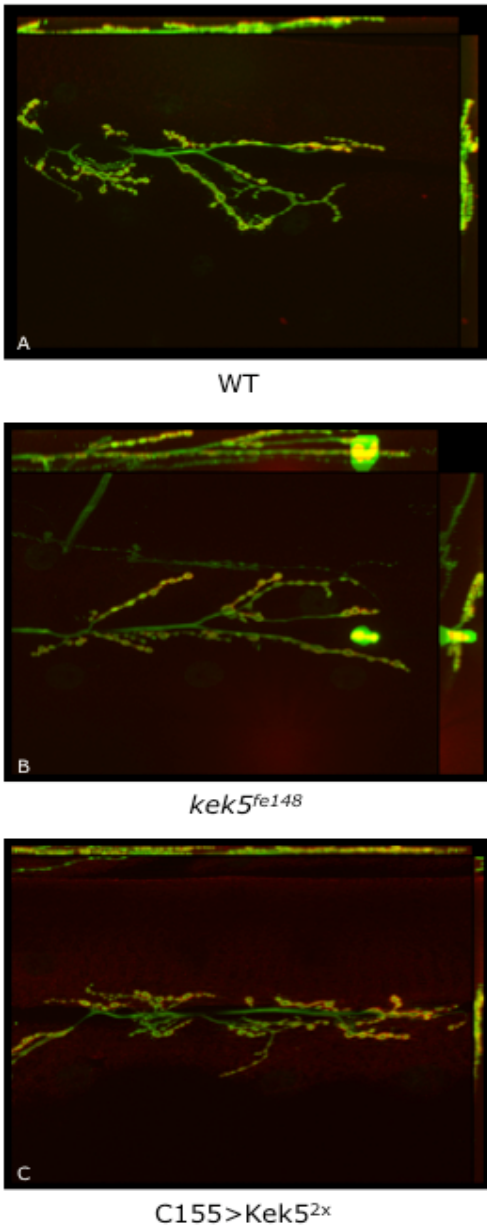


Figure 16: High resolution (40X) images of NMJ. Epifluorescent micrographs of the NMJ of muscle 6/7 of 3rd instar larva as in Fig. 14, z-stack MIP projection, apotome. processing.

DISCUSSION

It is imperative to take advantage of the emerging tools and techniques to keep research at the forefront of science. Small silencing RNA was first discovered 15 years ago, and our knowledge of RNA has boomed in the last decade. RNAi has since been the focus of much interest and press and promises to be a powerful and convenient technique to promote gene silencing in both research and therapeutic fronts.

With genome-wide transgenic RNAi libraries made available from the VDRC and similar projects (i.e. TRIP at Harvard), the *Drosophila* community has a new set of tools to advance genetic inquiries (Ni et al., 2009). Here I have validated the identity and knockdown effectiveness of several RNAi lines targeted against the members of the Kek family for current and future investigations of their function. Furthermore, effectiveness of RNAi mediated knockdown in the developing nervous system was assessed, making this information available for the first time. Although some knockdown is observed in the developing nervous system, this effect is unlikely to be significant enough to result in detectable phenotypes, particularly considering that neuronal phenotypes are often subtle.

In addition our data indicates that coexpression of *dcr2* greatly increases the effectiveness of RNAi mediated knockdown, promoting significant knockdown in the embryo and early larva, and virtually complete knockdown in late larval stages. This data suggests that Dcr2 processing is

the limiting step in the processing of RNAi in the nervous system; hence future neuronal studies using transgenic RNAi must incorporate UASDcr2 for effective results.

It should be noted that effectiveness of knockdown only of Kek5 RNAi was tested in the developing nervous system. Although all lines presented similar knockdown effectiveness in adult eye, it is possible that different lines may have different effectiveness temporally in the nervous system.

SCREEN FOR KEK FAMILY FUNCTION

Neural knockdown of the *keks* with *dcr* coexpression shows no effect on adult behavior with the exception of *kek1* knockdown, which based on preliminary analyses shows aspects of locomotor deficits. Observations from this screen also indicate possible interactions between *kek1* and *kek2*, *nrt* and *nrg.* Since no interaction between *kek1* and *nrg* has been published to date, this is a novel result worth further investigating. *kek1* has been previously shown to interact with *nrt* in axonal guidance and fasciculation as detected in the axonal pattern of the CNS. It would be of interest to investigate if *kek1*/*nrg* double knockdown would also generate a phenotype in the same tissue.

Since *kek2* was previously shown to interact in the NMJ structure (Guan et al., 2005), it would also be worthwhile investigating if *kek1*/*kek2* double knockdown shows a synergistic effect in the NMJ. Furthermore, since

various molecules such as FasII have been shown to interact both in axonal guidance and fasciculation as well as NMJ structure, the potential for an interaction of *kek1* with *nrg* and *nrt* should also be investigated in the context of NMJ structure.

KEK5 ROLE IN THE NMJ

BMP signaling regulates NMJ size and structure via retrograde signaling from the muscle to the pre-synaptic MN (Keshishian and Kim, 2004). An increase in BMP signaling promotes growth of the NMJ as can be observed by increased bouton numbers. Given the *kek5* inhibitory interaction with BMP signaling in wing pattern formation (Evans et al., 2009), it was proposed that *kek5* null mutant would increase NMJ size due to increased BMP signaling. Likewise, *kek5* misexpression would inhibit BMP signaling, thereby inhibiting NMJ growth.

Preliminary analysis of NMJ indicates that this is not the case; albeit no statistically significant effect was observed, decreased bouton numbers was seen in the loss of function, while no trend was detected in the *kek5* gain of function. Furthermore pre-synaptic knockdown of *kek5* shows increased bouton number, indicating that if these trends are real, *kek5* is involved in a different mechanism of NMJ regulation. NMJ analysis at a higher resolution should be carried out to support and extend preliminary results. In addition, manipulating levels of *kek5* post-synaptically may provide further insight

into any possible involvement of *kek5* in NMJ regulation. Similarly, although a possible synergistic interaction was observed between *kek5* and *kek6* in the wing cross vein formation, no interaction was detected in the NMJ structure, nor was any effect of *kek6* itself observed on the NMJ.

RNAi AS A SCREENING TOOL

The screening of the Kekkon family function produced some results that need to be reconciled with previous data, such as the effects on viability of *kek1* and *kek3* ubiquitous knockdown. A meticulous evaluation of single *kek1* and *kek3* null mutant viability may confirm this novel phenotype, and would also validate transgenic RNAi as a legitimate tool for phenotypic screens.

Another unexpected result was the cross vein defect observed with *kek6* knockdown in imaginal disc. A *kek6* mutant has been previously characterized in the lab and no wing phenotype was observed. However, it is not clear that the characterized mutant was indeed a null mutation. Another observation of concern is the variation of penetrance between different RNAi lines. Although position effects due to chromosomal location of the transgenes can be expected among lines, such drastic variation is unusual.

Much is still unknown about the pathways controlling RNAi and it is uncertain what other possible artifacts may arise from using this technique. New classes of small RNAs are constantly being discovered and many still

have unclear biosynthesis, functions and pathway components. Moreover, cross talk among the different small RNA pathways have recently been detected. This is particularly concerning given that many classes of small RNA regulate chromatin states and general transcription. Furthermore, various components interact in the processing of different types of small RNA, and expressing high levels of exogenous RNAi triggers may titrate these components from other endogenous functions.

Another serious concern regards the use of Dcr2 to promote knockdown in the developing nervous system. Although efficiency of knockdown is significantly increased, raising the levels of Dcr2 may have detrimental effects to the organism and to the nervous system itself. For instance, in addition to exogenous dsRNA, Dicer is also involved in endogenous small silencing RNA (Chung et al., 2008; Okamura et al., 2008) and small nucleolar RNAs (snoRNAs) (Taft et al., 2009), some of which are actively regulated. Altering Dcr2 levels may affect the regulation of other genes, with unpredictable outcomes. Hence RNAi can be used as a powerful tool to screen and identify relevant phenotypes, but careful confirmation and validation of phenotypes with conventional and cleaner genetic techniques must follow initial identification.

MATERIALS AND METHODS

GENETICS

Fly stocks were kept at room temperature (24°C) and raised on standard media. Experimental crosses involving the Gal4/UAS system were raised on 28°C on standard media. Flies used for NMJ analysis were raised at 28°C in brown media. w^{1118} was used as the wild type strain, $Kek5^{2x}$ is [UASKek5GFP][16, 52] and the $Kek5^{IC\Delta123}$ line used was CE1-12F-1M.

RNAi transgenic lines were obtained from Vienna *Drosophila* Research Center (VDRC) (table 1) and presence of each construct was verified by PCR using standard procedures and gene specific primers as indicated in appendix 2. *Drosophila* genomic DNA templates were prepared by squishing a single fly in 50μl a solution of 10mM Tris buffer pH8, 1mM EDTA, 2.5mM NaCl and 0.2mg/ml proteinase K. This was then incubated at room temperature for 20min., heated to 95°C in a hot block for 2min. and centrifuged at maximum speed for 5 minutes in a microfuge. 4-8μl of this genomic prep was then used for PCR reactions.

Recombinant lines were created according scheme in Fig17 and validated according to table2. Table 2 also includes line number used to identify different lines. Roman numerals in these recombinants lines are used to identify various copies (different recombination events) of the same genotype.

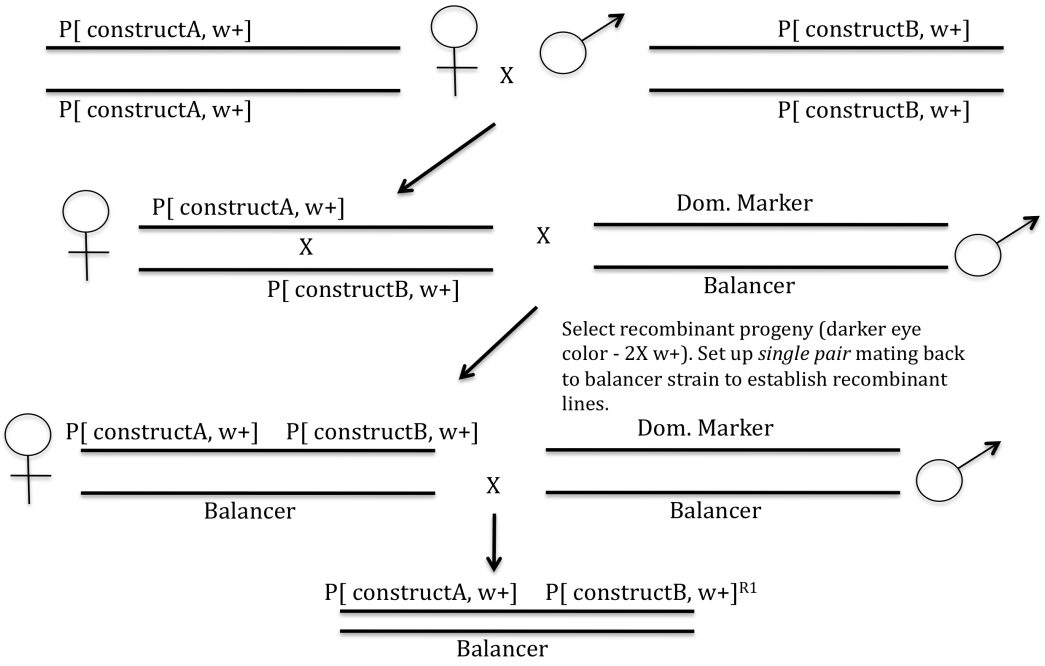


Figure 17: Schematic for creation of recombinants.

Recombinants lines used for validation of lines are listed below. GMR recombinants with Kek1 and Kek5 had been previously created Remaining recombinant lines were created by scheme in figure below, and selected by expression of GFP or PCR.

- GMRGAL4, [UAS.Kek1.GFP]^{59II}
- GMRGAL4, [UAS.Kek2.GFP]$^{R6\text{-}1a}$
- GMRGAL4, [UAS.Kek4.GFP]$^{7\text{-}4,\ 38\text{-}5}$
- GMRGAL4, [UAS.Kek5.GFP]$^{16,\ 52}$
- GMRGAL4, [UAS.Kek6.GFP]^{12}
- C155GAL4,UASkek5$^{RNAi\text{-}47770}$
- [UAS.kek5.GFP]^{11} ,UASkek5$^{RNAi\text{-}27249}$

Driver lines used are listed below with a Stock number. Full genotype and further information can be found at Flybase or Bloomington Drosophila Stock Center:

- GMRGAL4 - 1104
- apGAL4 - 3041
- tubGAL4 - 5138
- actGAL4 - 4414
- actGAL4; UASdcr-2 - 25708
- ElavGAL4 - 8765
- C155GAL4 - 458
- C155GAL4; UASdcr-2 -25750
- enGAL4 - 6356
- ScabrousGAL4 -6479

ASSESSMENT OF KNOCKDOWN IN DEVELOPING NERVOUS SYSTEM

The lines used for RNAi knockdown assessment in the developing nervous system were *C155GAL4;UASDcr2*^{25750}, [*C155GAL4*,*UASkek5*^{RNAi}]^{III}, [*UASkek5-GFP*,*UASkek5*^{RNAi}]^{V} and *UASkek5-GFP*^{16}.

Embryos – Live embryos were washed into collection nets, placed into petri dishes, dechorionated using 50% bleach. The bleach was removed and vitelline membranes were then removed by popping embryos with a sharp tungstein needle in water after adhering embryos to the petri dish plate by allowing them to dry briefly. Embryos were then mounted in water and imaged immediately.

1st instar larva – live larva 24-30hrs after egg lay were incubated in PBT with 4% formaldehyde for 15-20 min, then mounted in water and imaged immediately.

3rd instar larva – wandering larval brains were dissected and fixed in PBT with 4% formaldehyde for 15 min, then mounted in mounting media (50% glycerol in PBS).

IMMUNOHISTOCHEMISTRY - EMBRYOS

Embryos were collected overnight, dechorionated with 50% bleach and fixed for 20 minutes in a mixture of 6mls heptane and 1ml fixative (10% formaldehyde and 50mM EGTA in PBS). Fixative was then removed (bottom layer), 8mls of methanol added and solution was shaken vigorously for 1 min. Embryos were then extensively rinsed with 100% methanol and stored in methanol at -20°C. Embryos were rehydrated in PBS with 0.1% Tween20 and blocked for 30minutes in PBT with 5% NGS at room temperature. Primary antibody was incubated overnight at 4°C in PBT with 5% normal goat serum (NGS). Primary antibodies and dilutions were as follows: anti-GFP (BD Bioscience) 1:500, BP102 (DSHB) 1:500, and 1D4 (DSHB) 1:100. Secondary antibody was incubated for two hours in room temperature in PBT with 5% normal goat serum. The secondary antibody was anti-mouse or rabbit Alexa 488 used at 1:500. Samples were mounted in 70% Glycerol in PBS.

IMMUNOHISTOCHEMISTRY - LARVA

Dissecting magnetic chambers were constructed similar to published protocol (Bellen & Budnik, 2000), with the necessary adaptations as indicated in Appendix 1. Third instar wandering larva were selected and dissected in Ca^{+} free saline (128mM NaCl, 2mMKCl, 4mM $MgCl_2$, 35.5mM Sucrose, 5mM Hepes, 1mM EGTA). Larvae were pinned down by posterior and anterior extremities and opened dorsally along midline using dissecting scissors (Roboz RS5618). Interior organs and gut were removed with fine forceps and larval cuticle (with attached muscles) were opened and pinned at the 4 extremities so that cuticle of larva is flat against chamber surface. While pinned down, larvae were fixed with 4% formaldehyde (alcohol free) in 0.1M phosphate buffer pH7.2. Samples were then transferred to round bottom wells (maximum 4/well), washed and stained according to standard immunostaining procedures using 0.1M phosphate buffer pH7.2 with 0.2% TritonX-100. Samples were arranged in desired orientation and mounted cuticle downward in glycerol/PBS. Primary antibodies used were rabbit anti-HRP (Jackson ImmunoResearch) – 1:1000, mouse anti-Discs large (DSHB) – 1:2000 and rabbit Anti-GFP – 1:1000. Secondary antibodies used were goat anti mouse and rabbit Alexa 488 or 568 at 1:500.

REFERENCES

Abelson, J. F., Kwan, K. Y., O'Roak, B. J., Baek, D. Y., Stillman, A. A., Morgan, T. M., Mathews, C. A., Pauls, D. L., Rasin, M. R., Gunel, M. et al. (2005). Sequence variants in SLITRK1 are associated with Tourette's syndrome. *Science* **310**, 317-20.

Aberle, H., Haghighi, A. P., Fetter, R. D., McCabe, B. D., Magalhaes, T. R. and Goodman, C. S. (2002). wishful thinking encodes a BMP type II receptor that regulates synaptic growth in Drosophila. *Neuron* **33**, 545-58.

Alvarado, D., Rice, A. H. and Duffy, J. B. (2004). Knockouts of Kekkon1 define sequence elements essential for Drosophila epidermal growth factor receptor inhibition. *Genetics* **166**, 201-11.

Ashley, J., Packard, M., Ataman, B. and Budnik, V. (2005). Fasciclin II signals new synapse formation through amyloid precursor protein and the scaffolding protein dX11/Mint. *J Neurosci* **25**, 5943-55.

Bando, T., Sekine, K., Kobayashi, S., Watabe, A. M., Rump, A., Tanaka, M., Suda, Y., Kato, S., Morikawa, Y., Manabe, T. et al. (2005). Neuronal leucine-rich repeat protein 4 functions in hippocampus-dependent long-lasting memory. *Mol Cell Biol* **25**, 4166-75.

Baum, P. D. and Garriga, G. (1997). Neuronal migrations and axon fasciculation are disrupted in ina-1 integrin mutants. *Neuron* **19**, 51-62.

Bearer, C. F. (2001). Mechanisms of brain injury: L1 cell adhesion molecule as a target for ethanol-induced prenatal brain injury. *Semin Pediatr Neurol* **8**, 100-7.

Bech-Hansen, N. T., Naylor, M. J., Maybaum, T. A., Sparkes, R. L., Koop, B., Birch, D. G., Bergen, A. A., Prinsen, C. F., Polomeno, R. C., Gal, A. et al. (2000). Mutations in NYX, encoding the leucine-rich

proteoglycan nyctalopin, cause X-linked complete congenital stationary night blindness. *Nat Genet* **26**, 319-23.

Bieber, A. J., Snow, P. M., Hortsch, M., Patel, N. H., Jacobs, J. R., Traquina, Z. R., Schilling, J. and Goodman, C. S. (1989). Drosophila neuroglian: a member of the immunoglobulin superfamily with extensive homology to the vertebrate neural adhesion molecule L1. *Cell* **59**, 447-60.

Budnik, V. (1996). Synapse maturation and structural plasticity at Drosophila neuromuscular junctions. *Curr Opin Neurobiol* **6**, 858-67.

Campos-Ortega, J. A. (1993). Early Neurogenesis in Drosophila Melanogaster. In *The Develoment of Drosophila melanogaster*, vol. II (ed. A. M. A. M. Bate): Cold Spring Harbor Lab Press.

Campos-Ortega, J. A. (1995). Genetic mechanisms of early neurogenesis in Drosophila melanogaster. *Mol Neurobiol* **10**, 75-89.

Campos-Ortega, J. A. H., V. (1997). The embryonic development of *Drosophila melanogaster*. Germany: Springer.

Chen, Y., Aulia, S., Li, L. and Tang, B. L. (2006). AMIGO and friends: an emerging family of brain-enriched, neuronal growth modulating, type I transmembrane proteins with leucine-rich repeats (LRR) and cell adhesion molecule motifs. *Brain Res Rev* **51**, 265-74.

Chilton, J. K. (2006). Molecular mechanisms of axon guidance. *Dev Biol* **292**, 13-24.

Chung, W. J., Okamura, K., Martin, R. and Lai, E. C. (2008). Endogenous RNA interference provides a somatic defense against Drosophila transposons. *Curr Biol* **18**, 795-802.

Cunningham, B. A., Hemperly, J. J., Murray, B. A., Prediger, E. A., Brackenbury, R. and Edelman, G. M. (1987). Neural cell adhesion molecule: structure, immunoglobulin-like domains, cell surface modulation, and alternative RNA splicing. *Science* **236**, 799-806.

Dietzl, G., Chen, D., Schnorrer, F., Su, K. C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblauer, S. et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature* **448**, 151-6.

Duffy, J. B. (2002). GAL4 system in Drosophila: a fly geneticist's Swiss army knife. *Genesis* **34**, 1-15.

Duffy, J. B. and Gergen, J. P. (1994). Sex, segments, and the CNS: Common mechanisms used to determine cell fates in *Drosophila*. *Advances in Genetics* **31**, 1-28.

Evans, T. A., Haridas, H. and Duffy, J. B. (2009). Kekkon5 is an extracellular regulator of BMP signaling. *Dev Biol* **326**, 36-46.

Fransen, E., Vits, L., Van Camp, G. and Willems, P. J. (1996). The clinical spectrum of mutations in L1, a neuronal cell adhesion molecule. *Am J Med Genet* **64**, 73-7.

Ghiglione, C., Carraway, K. L., 3rd, Amundadottir, L. T., Boswell, R. E., Perrimon, N. and Duffy, J. B. (1999). The transmembrane molecule kekkon 1 acts in a feedback loop to negatively regulate the activity of the Drosophila EGF receptor during oogenesis. *Cell* **96**, 847-56.

Ghildiyal, M. and Zamore, P. D. (2009). Small silencing RNAs: an expanding universe. *Nat Rev Genet* **10**, 94-108.

Goodman, C. S. and Doe, C. Q. (1993). Embryonic development of the *Drosophila* central nervous system. In *The Development of Drosophila melanogaster*, vol. II (ed. M. A. Bate, A.M.): Cold Spring Harbor Lab. Press.

Goold, C. P. and Davis, G. W. (2007). The BMP ligand Gbb gates the expression of synaptic homeostasis independent of synaptic growth control. *Neuron* **56**, 109-23.

Guan, Z., Saraswati, S., Adolfsen, B. and Littleton, J. T. (2005). Genome-wide transcriptional changes associated with enhanced activity in the Drosophila nervous system. *Neuron* **48**, 91-107.

Hall, S. G. and Bieber, A. J. (1997). Mutations in the Drosophila neuroglian cell adhesion molecule affect motor neuron pathfinding and peripheral nervous system patterning. *J Neurobiol* **32**, 325-40.

Hoang, B. and Chiba, A. (1999). 'Identify' and 'lock in': molecular integration during synaptic target recognition. *Cell Mol Life Sci* **55**, 1399-406.

Hoang, B. and Chiba, A. (2001). Single-cell analysis of Drosophila larval neuromuscular synapses. *Dev Biol* **229**, 55-70.

Ishii, N., Wanaka, A. and Tohyama, M. (1996). Increased expression of NLRR-3 mRNA after cortical brain injury in mouse. *Brain Res Mol Brain Res* **40**, 148-52.

Iwai, Y., Usui, T., Hirano, S., Steward, R., Takeichi, M. and Uemura, T. (1997). Axon patterning requires DN-cadherin, a novel neuronal adhesion receptor, in the Drosophila embryonic CNS. *Neuron* **19**, 77-89.

Kalachikov, S., Evgrafov, O., Ross, B., Winawer, M., Barker-Cummings, C., Martinelli Boneschi, F., Choi, C., Morozov, P., Das, K., Teplitskaya, E. et al. (2002). Mutations in LGI1 cause autosomal-dominant partial epilepsy with auditory features. *Nat Genet* **30**, 335-41.

Katidou, M., Vidaki, M., Strigini, M. and Karagogeos, D. (2008). The immunoglobulin superfamily of neuronal cell adhesion molecules: lessons from animal models and correlation with human disease. *Biotechnol J* **3**, 1564-80.

Keshishian, H. and Kim, Y. S. (2004). Orchestrating development and function: retrograde BMP signaling in the Drosophila nervous system. *Trends Neurosci* **27**, 143-7.

Klementiev, B., Novikova, T., Novitskaya, V., Walmod, P. S., Dmytriyeva, O., Pakkenberg, B., Berezin, V. and Bock, E. (2007). A neural cell adhesion molecule-derived peptide reduces neuropathological signs and cognitive impairment induced by Abeta25-35. *Neuroscience* **145**, 209-24.

Lin, D. M., Fetter, R. D., Kopczynski, C., Grenningloh, G. and Goodman, C. S. (1994). Genetic analysis of Fasciclin II in Drosophila: defasciculation, refasciculation, and altered fasciculation. *Neuron* **13**, 1055-69.

Lin, D. M. and Goodman, C. S. (1994). Ectopic and increased expression of Fasciclin II alters motoneuron growth cone guidance. *Neuron* **13**, 507-23.

Lin, J. C., Ho, W. H., Gurney, A. and Rosenthal, A. (2003). The netrin-G1 ligand NGL-1 promotes the outgrowth of thalamocortical axons. *Nat Neurosci* **6**, 1270-6.

MacLaren, C. M., Evans, T. A., Alvarado, D. and Duffy, J. B. (2004). Comparative analysis of the Kekkon molecules, related members of the LIG superfamily. *Dev Genes Evol* **214**, 360-6.

Maness, P. F. and Schachner, M. (2007). Neural recognition molecules of the immunoglobulin superfamily: signaling transducers of axon guidance and neuronal migration. *Nat Neurosci* **10**, 19-26.

McCabe, B. D., Hom, S., Aberle, H., Fetter, R. D., Marques, G., Haerry, T. E., Wan, H., O'Connor, M. B., Goodman, C. S. and Haghighi, A. P. (2004). Highwire regulates presynaptic BMP signaling essential for synaptic growth. *Neuron* **41**, 891-905.

Miller, C. M., Page-McCaw, A. and Broihier, H. T. (2008). Matrix metalloproteinases promote motor axon fasciculation in the Drosophila embryo. *Development* **135**, 95-109.

Musacchio, M. and Perrimon, N. (1996). The Drosophila kekkon genes: novel members of both the leucine-rich repeat and immunoglobulin superfamilies expressed in the CNS. *Dev Biol* **178**, 63-76.

Nakayama, H., Kazama, H., Nose, A. and Morimoto-Tanifuji, T. (2006). Activity-dependent regulation of synaptic size in Drosophila neuromuscular junctions. *J Neurobiol* **66**, 929-39.

Ni, J. Q., Liu, L. P., Binari, R., Hardy, R., Shim, H. S., Cavallaro, A., Booker, M., Pfeiffer, B., Markstein, M., Wang, H. et al. (2009). A Drosophila Resource of Transgenic RNAi Lines for Neurogenetics. *Genetics*.

Okamura, K., Balla, S., Martin, R., Liu, N. and Lai, E. C. (2008). Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster. *Nat Struct Mol Biol* **15**, 998.

Ono, T., Sekino-Suzuki, N., Kikkawa, Y., Yonekawa, H. and Kawashima, S. (2003). Alivin 1, a novel neuronal activity-dependent gene, inhibits apoptosis and promotes survival of cerebellar granule neurons. *J Neurosci* **23**, 5887-96.

Prokop, A. and Meinertzhagen, I. A. (2006). Development and structure of synaptic contacts in Drosophila. *Semin Cell Dev Biol* **17**, 20-30.

Pusch, C. M., Zeitz, C., Brandau, O., Pesch, K., Achatz, H., Feil, S., Scharfe, C., Maurer, J., Jacobi, F. K., Pinckers, A. et al. (2000). The complete form of X-linked congenital stationary night blindness is caused by mutations in a gene encoding a leucine-rich repeat protein. *Nat Genet* **26**, 324-7.

Ritzenthaler, S. and Chiba, A. (2001). Multiple personalities: synaptic target cells as introverts and extroverts. *Dev Growth Differ* **43**, 503-8.

Rose, D. and Chiba, A. (1999). A single growth cone is capable of integrating simultaneously presented and functionally distinct molecular cues during target recognition. *J Neurosci* **19**, 4899-906.

Rose, D. and Chiba, A. (2000). Synaptic target recognition at Drosophila neuromuscular junctions. *Microsc Res Tech* **49**, 3-13.

Rose, D., Zhu, X., Kose, H., Hoang, B., Cho, J. and Chiba, A. (1997). Toll, a muscle cell surface molecule, locally inhibits synaptic initiation of the RP3 motoneuron growth cone in Drosophila. *Development* **124**, 1561-71.

Sanchez-Soriano, N., Tear, G., Whitington, P. and Prokop, A. (2007). Drosophila as a genetic and cellular model for studies on axonal growth. *Neural Dev* **2**, 9.

Siller, K. H. and Doe, C. Q. (2009). Spindle orientation during asymmetric cell division. *Nat Cell Biol* **11**, 365-74.

Speicher, S., Garcia-Alonso, L., Carmena, A., Martin-Bermudo, M. D., de la Escalera, S. and Jimenez, F. (1998). Neurotactin functions in concert with other identified CAMs in growth cone guidance in Drosophila. *Neuron* **20**, 221-33.

Taft, R. J., Glazov, E. A., Lassmann, T., Hayashizaki, Y., Carninci, P. and Mattick, J. S. (2009). Small RNAs derived from snoRNAs. *RNA* **15**, 1233-40.

Tsuji, L., Yamashita, T., Kubo, T., Madura, T., Tanaka, H., Hosokawa, K. and Tohyama, M. (2004). FLRT3, a cell surface molecule containing LRR repeats and a FNIII domain, promotes neurite outgrowth. *Biochem Biophys Res Commun* **313**, 1086-91.

Van Vactor, D. (1998). Adhesion and signaling in axonal fasciculation. *Curr Opin Neurobiol* **8**, 80-6.

APPENDIX

APPENDIX 1

MAGNETIC CHAMBERS AND PINS - ASSEMBLY

Materials

- Glass slides – Glass cut into 2''x3'', approx 1/16-1/8'' thickness obtained from a window store. Edges were polished for safety.
- Adhesive magnet business card size obtained from office supply store.
- Silicone sealant
- Beaded stainless steel insect pins size 00 obtained from Carolina Biological Supply Co. item#65-4331
- Prong fasteners from office supply store
- Epoxy glue (5 min.)
- Fine thin-nose pliers
- Sharp cutting pliers
- Snap blade knife
- Sand paper grit 200 and 400 (silicone carbide recommended).

Chambers Assembly

1) Using the adhesive in the magnetic sheet, glue 2 magnets sheets together.

2) Cut a hole around a nickel coin as a template with a snap blade knife.

4) Remove the adhesive protector and place a bead of silicone sealant around the perimeter of the strip and then again right at the rim of the cut hole (see diagram). Make sure there are no breaks around the inner rim.

5) Place the 2" x 3” glass slide over the back side of the magnetic strip and press into the adhesive. Some sealant should leak into the hole. Use the excess sealant to seal the inner edge of the magnetic assembly, to make sure no liquid will penetrate in between the magnetic sheets. Carefully place a heavy, flat object over the chambers for several hours (12-24hrs). A heavy book lined with aluminum foil or wax paper works well. Some sealants release acetic acid, so fill chambers with water and let sit 12-24hrs.

Pins assembly

1) Form beaded end insect pins as in Fig. 1 by using a thin-nosed pair of pliers. Rinse the formed end with 95% alcohol.

2) Cut prong fastener approx. 3-5mm before fold (the thinner side) and about 15mm after fold (bottom part), as indicated in the diagram above figures. Using pliers fold handle end in the middle, then fold again the edges 90° to have it stand up on its own. For safety and comfort, also fold raised edge from the tab base. Final assembly should look like figure 2.

3) Sand down the surfaces to be glued. This makes epoxy adhere better. Clean and rinse with alcohol to remove any oils from fingers, etc.

4) Using quick setting epoxy, glue pins to tabs as in Fig. 2. Apply one drop of glue from a toothpick or wooden applicator stick. Middle section left from prong fasteners works well. The pin tip should be raised about 2-3mm (set tip on a strip of cardboard, etc. before gluing, diagram on bottom of figures). This raised angle will prevent the solution from

spreading all over the top of the chamber via capillary action. Let glue sit overnight to harden completely.

5) After glue has hardened, cut the two center pins to ~16 mm and the 4 corner pins to ~18-20 mm (for 2^{nd}-3^{rd} instars) , measured from the end of the tab (the end with the glue). Pin tips should be sanded flat and thin with silicon carbide paper - grit #220. Finish with grit # 400-600. See diagram. Tip can be sharp but not too narrow. It's convenient for your index fingernail to have 2 mm of white showing.

5) Bend the pins as in Fig. 3. Dimensions are in millimeters. The last bend (0.5-1mm) is best done with old and somewhat dull forceps under a low power dissecting scope.

6) Corner pins should also be bent from the side as in Fig. 4. Pin tips should have moderate to light tension when bent to position on surface of glass. Too much tension rips cuticle, especially when larvae are very young. Not enough tension will not keep larva in place. All dimensions are approximate and may need to be adjusted for chamber.

Tab assembly

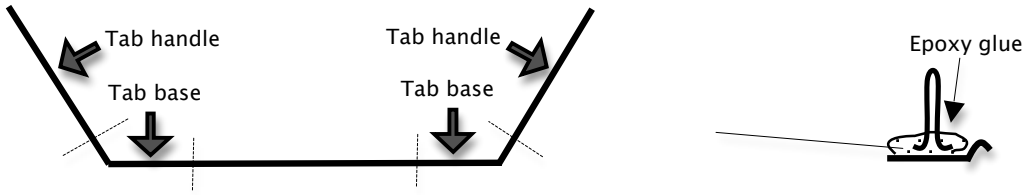


Fig. 1

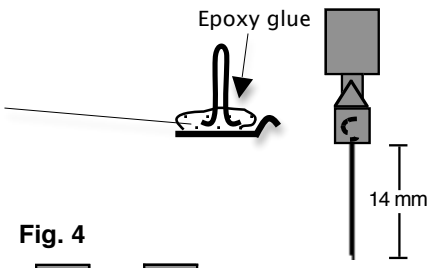
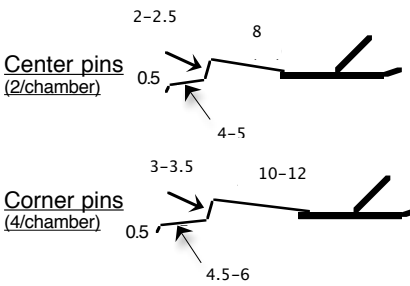


Fig. 2

Fig. 3

Length is measured from end of tab. See Fig. 2, bottom right.



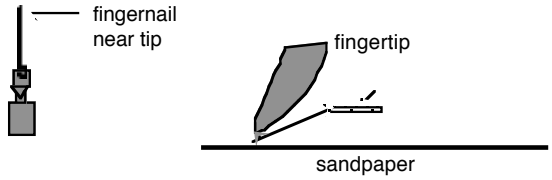
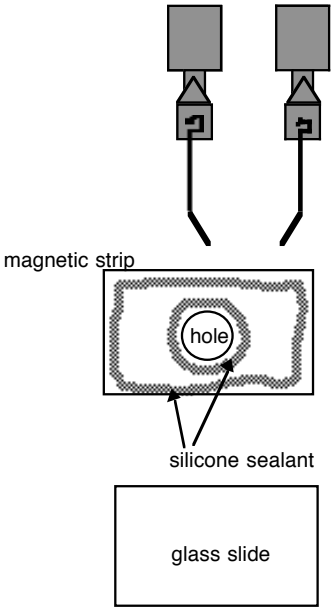


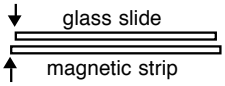
Fig. 4



8" x 1" strip of 220 silicone carbide paper; 2-3 passes, top and bottom of pin; get tip quite thin

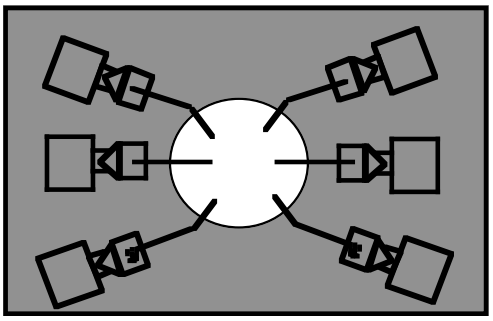
Then do likewise with 400-600 grit paper

Place glass slide over strip and position carefully to allow proper 1mm overlap of strip.

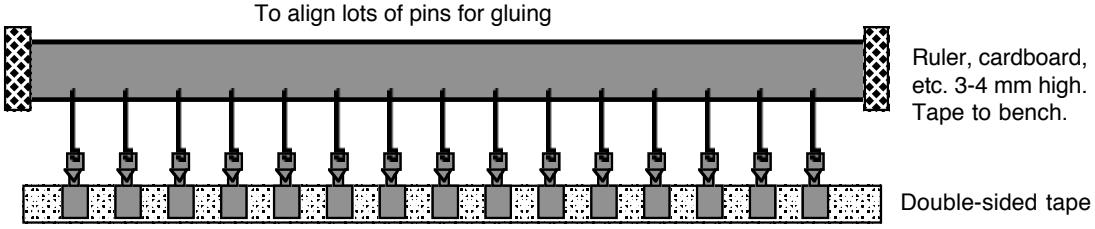


Invert above assembly and bend pins to provide the correct amount of pressure on slide.

Finished Chamber



To align lots of pins for gluing



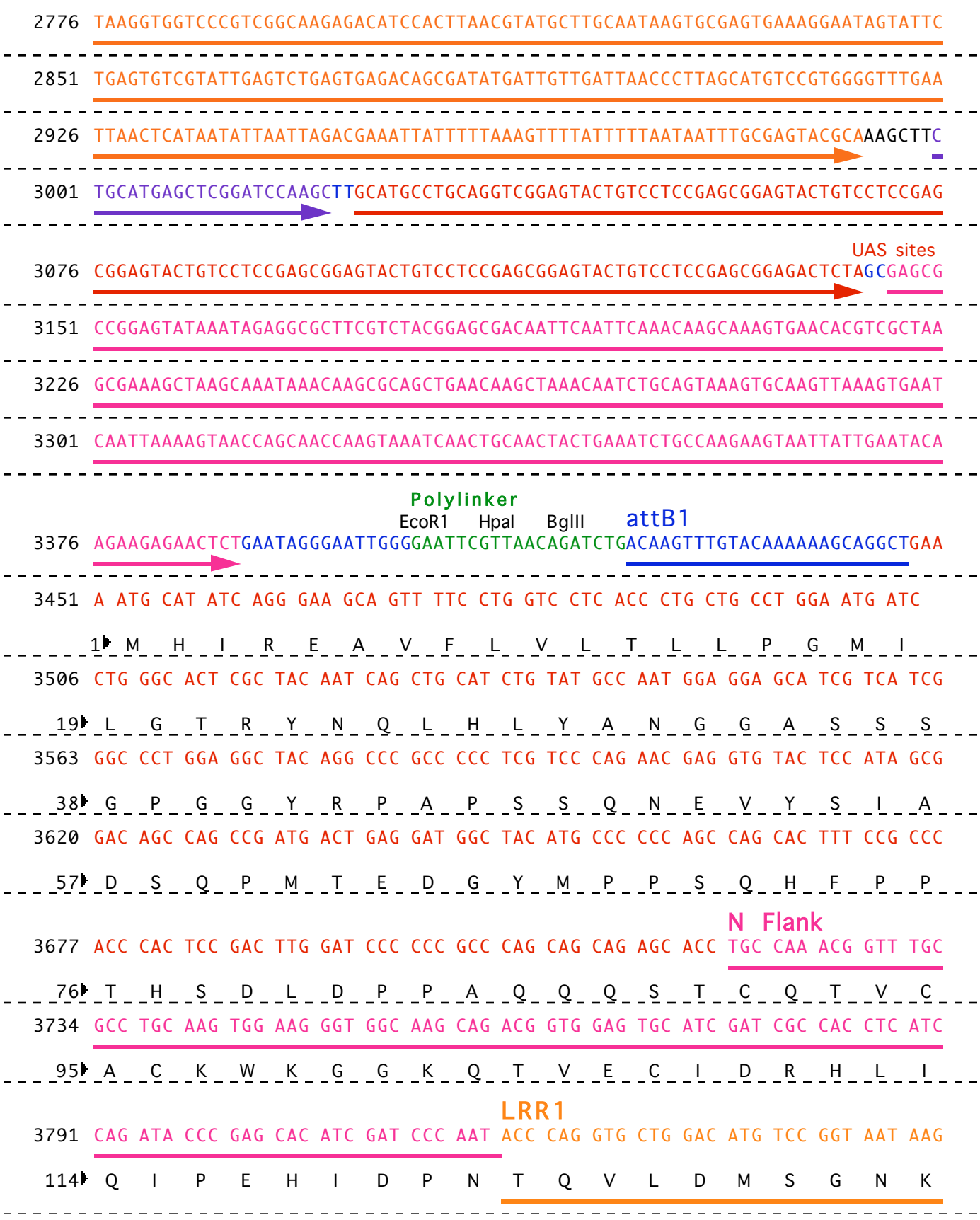
APPENDIX 2

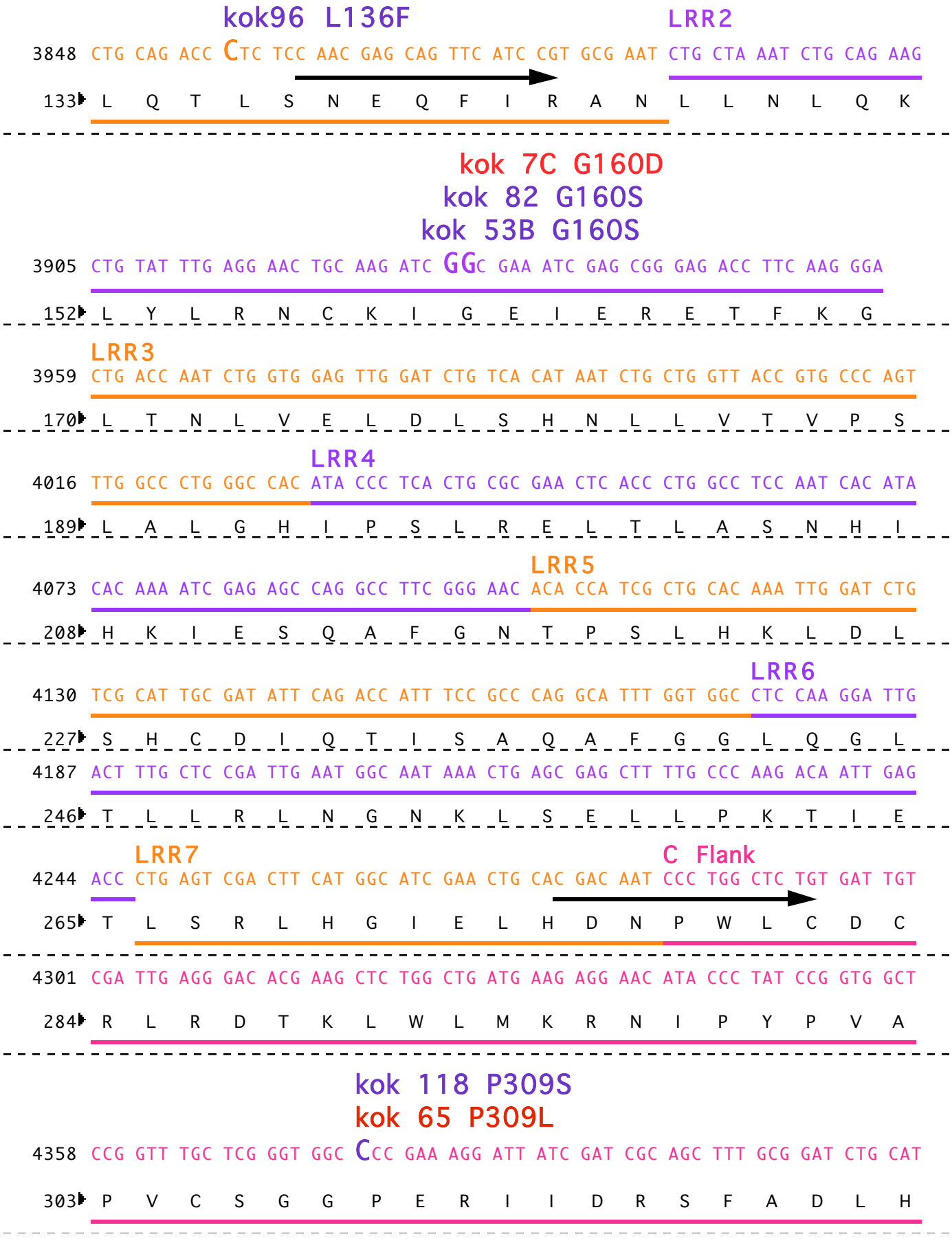
GENETIC SEQUENCES

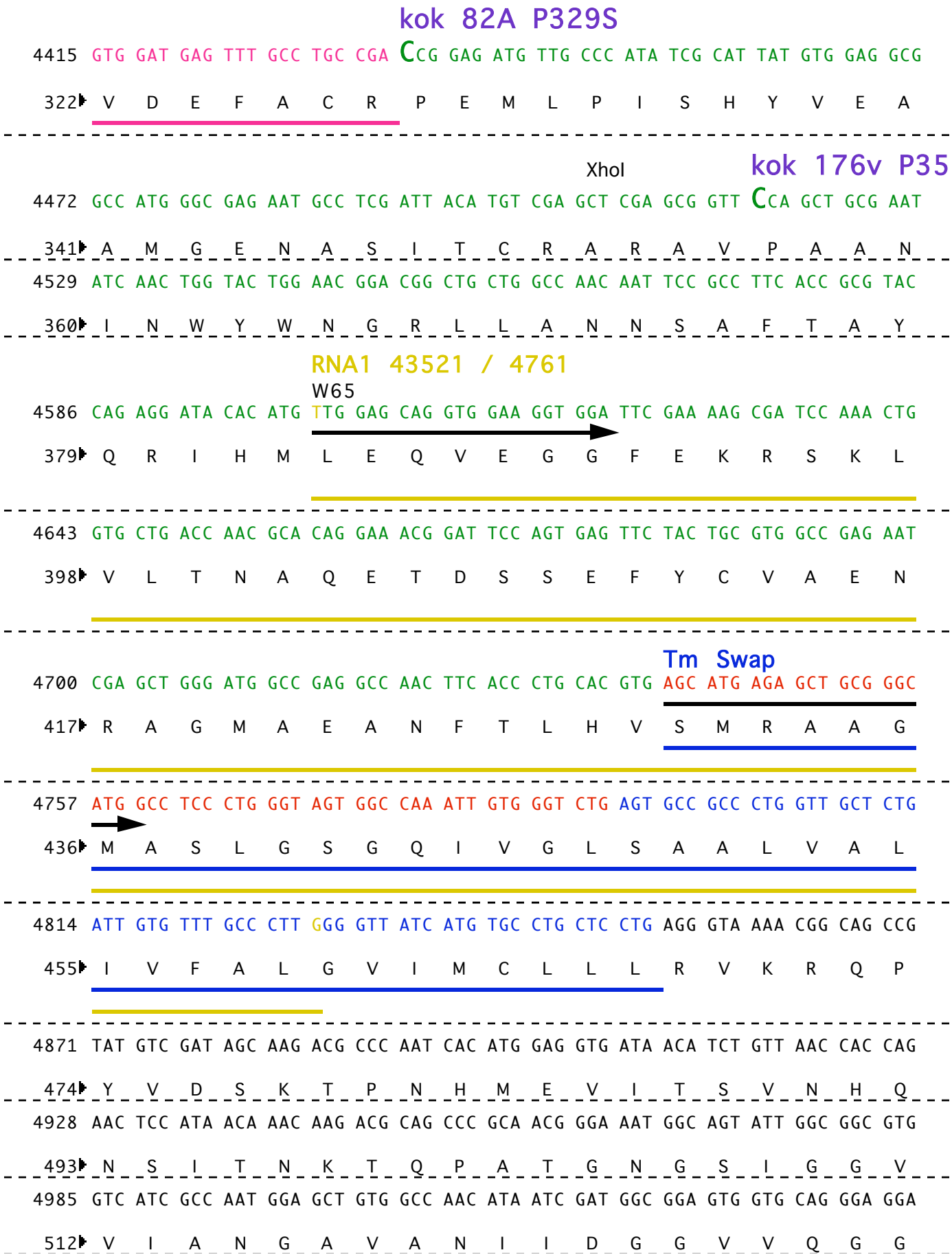
The sequences of the vectors used to clone individual keks are indicated here. Within each sequence is the open reading frame (ORF) of each kek, and within it, some structures are marked (such as LRRs and transmembrane region). Also marked (in dark yellow) are the trigger sequence of the RNAi lines, labeled as RNAi-#### (line number), and the oligo used in PCR to identify presence of construct, marked in black, labeled with oligo number starting with W (e.g. W65). RNai trigger and oligos are also marked on ORF of *nrt* and *nrg*.

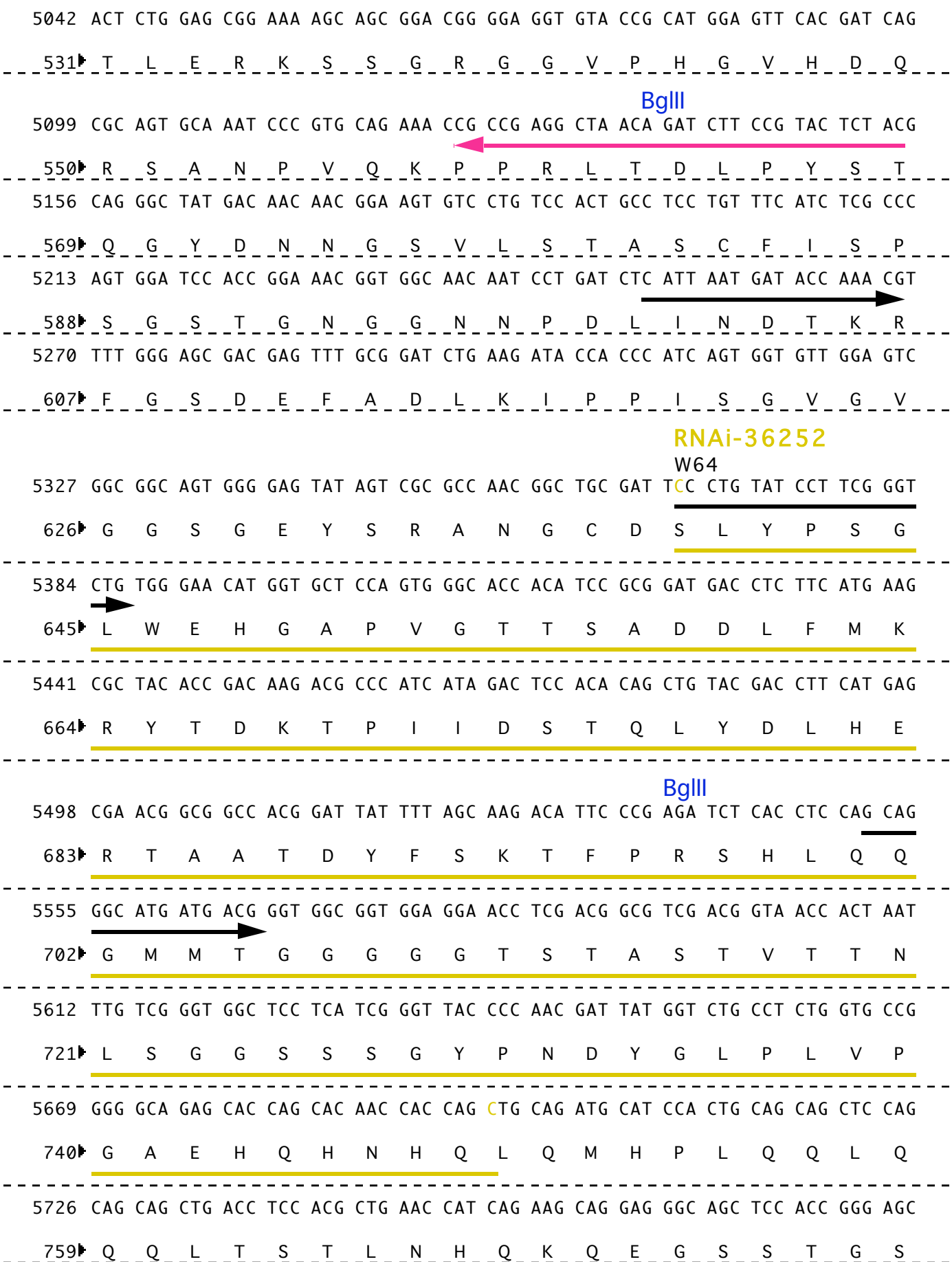
UAS Kek1 GFP

```
   1 GGCCAGACCCACGTAGTCCAGCGGCAGATCGGCGGCGGAGAAGTTAAGCGTCTCCAGGATGACCTTGCCCGAACT
  76 GGGGCACGTGGTGTTCGACGATGTGCAGCTAATTTCGCCCGGCTCCACGTCCGCCCATTGGTTAATCAGCAGACC
 151 CTCGTTGGCGTAACGGAACCATGAGAGGTACGACAACCATTTGAGGTATACTGGCACCGAGCCCGAGTTCAAGAA
 226 GAAGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA
 301 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC
 376 GCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCT
 451 CAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTT
 526 ATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG
 601 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAG
 676 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA
 751 ACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGA
 826 AGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG
 901 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGA
 976 GTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATC
1051 CATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAAT
1126 GATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG
1201 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCC
1276 AGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTC
1351 ATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTT
1426 CGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC
1501 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG
1576 TATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGT
1651 GCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTA
1726 ACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG
1801 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTA
1876 TTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGG
1951 GGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAA
2026 AAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCT
2101 CCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGT
2176 TGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGA
2251 AATACCGCACCGAATCGCGCGGAACTAACGACAGTCGCTCCAAGGTCGTCGAACAAAAGGTGAATGTGTTGCGGA
2326 GAGCGGGTGGGAGACAGCGAAAGAGCAACTACGAAACGTGGTGTGGTGGAGGTGAATTATGAAGAGGGCGCGCGA
2401 TTTGAAAAGTATGTATATAAAAAATATATCCCGGTGTTTTATGTAGCGATAAACGAGTTTTTGATGTAAGGTATG
2476 CAGGTGTGTAAGTCTTTTGGTTAGAAGACAAATCCAAAGTCTACTTGTGGGGATGTTCGAAGGGGAAATACTTGT
2551 ATTCTATAGGTCATATCTTGTTTTTATTGGCACAAATATAATTACATTAGCTTTTTGAGGGGGCAATAAACAGTA
2626 AACACGATGGTAATAATGGTAAAAAAAAAAACAAGCAGTTATTTCGGATATATGTCGGCTACTCCTTGCGTCGGG

                                                                        3'P
2701 CCCGAAGTCTTAGAGCCAGATATGCGAGCACCCGGAAGCTCACGATGAGAATGGCCAGACCATGATGAAATAACA
```

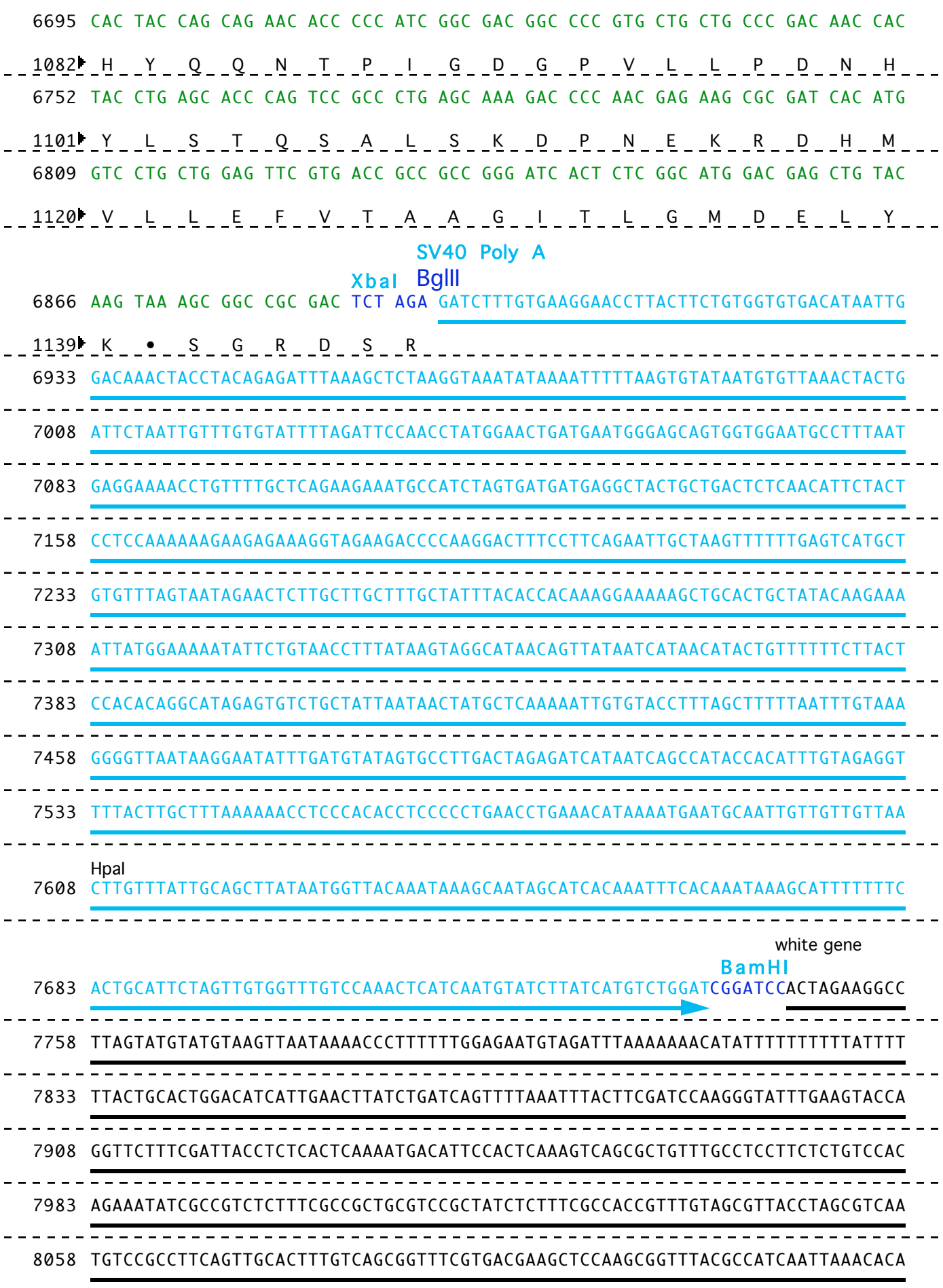

2776 TAAGGTGGTCCCGTCGGCAAGAGACATCCACTTAACGTATGCTTGCAATAAGTGCGAGTGAAAGGAATAGTATTC
2851 TGAGTGTCGTATTGAGTCTGAGTGAGACAGCGATATGATTGTTGATTAACCCTTAGCATGTCCGTGGGGTTTGAA
2926 TTAACTCATAATATTAATTAGACGAAATTATTTTTAAAGTTTTATTTTTAATAATTTGCGAGTACGCAAAGCTTC
3001 TGCATGAGCTCGGATCCAAGCTTGCATGCCTGCAGGTCGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAG
UAS sites
3076 CGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAGCGGAGACTCTAGCGAGCG
3151 CCGGAGTATAAATAGAGGCGCTTCGTCTACGGAGCGACAATTCAATTCAAACAAGCAAAGTGAACACGTCGCTAA
3226 GCGAAAGCTAAGCAAATAAACAAGCGCAGCTGAACAAGCTAAACAATCTGCAGTAAAGTGCAAGTTAAAGTGAAT
3301 CAATTAAAAGTAACCAGCAACCAAGTAAATCAACTGCAACTACTGAAATCTGCCAAGAAGTAATTATTGAATACA
Polylinker
EcoR1 HpaI BglII
attB1
3376 AGAAGAGAACTCTGAATAGGGAATTGGGGAATTCGTTAACAGATCTGACAAGTTTGTACAAAAAAGCAGGCTGAA
3451 A ATG CAT ATC AGG GAA GCA GTT TTC CTG GTC CTC ACC CTG CTG CCT GGA ATG ATC
1 M H I R E A V F L V L T L L P G M I
3506 CTG GGC ACT CGC TAC AAT CAG CTG CAT CTG TAT GCC AAT GGA GGA GCA TCG TCA TCG
19 L G T R Y N Q L H L Y A N G G A S S S
3563 GGC CCT GGA GGC TAC AGG CCC GCC CCC TCG TCC CAG AAC GAG GTG TAC TCC ATA GCG
38 G P G G Y R P A P S S Q N E V Y S I A
3620 GAC AGC CAG CCG ATG ACT GAG GAT GGC TAC ATG CCC CCC AGC CAG CAC TTT CCG CCC
57 D S Q P M T E D G Y M P P S Q H F P P
N Flank
3677 ACC CAC TCC GAC TTG GAT CCC CCC GCC CAG CAG CAG AGC ACC TGC CAA ACG GTT TGC
76 T H S D L D P P A Q Q Q S T C Q T V C
3734 GCC TGC AAG TGG AAG GGT GGC AAG CAG ACG GTG GAG TGC ATC GAT CGC CAC CTC ATC
95 A C K W K G G K Q T V E C I D R H L I
LRR1
3791 CAG ATA CCC GAG CAC ATC GAT CCC AAT ACC CAG GTG CTG GAC ATG TCC GGT AAT AAG
114 Q I P E H I D P N T Q V L D M S G N K

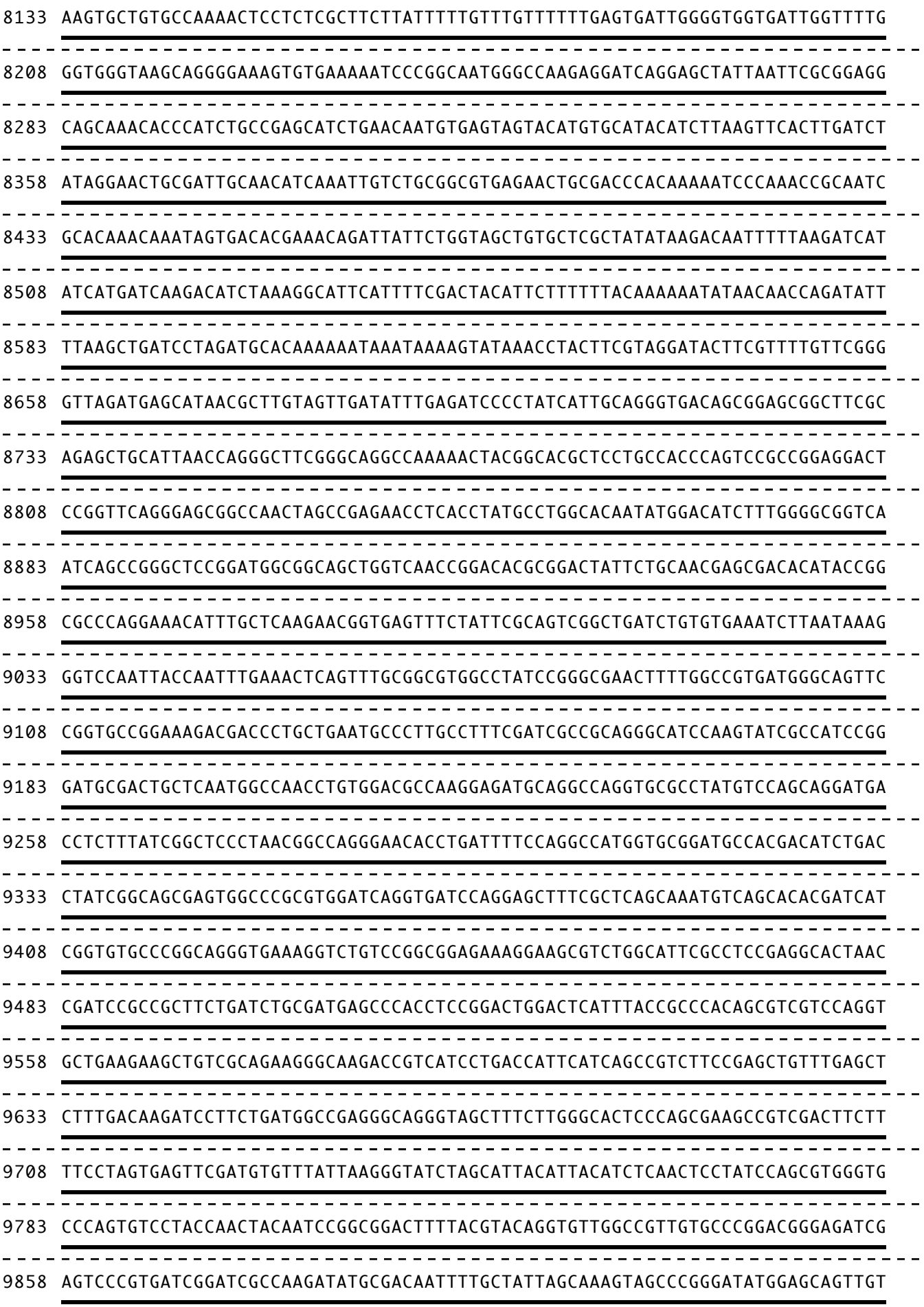

kok96 L136F
LRR2
3848 CTG CAG ACC CTC TCC AAC GAG CAG TTC ATC CGT GCG AAT CTG CTA AAT CTG CAG AAG
133 L Q T L S N E Q F I R A N L L N L Q K
kok 7C G160D
kok 82 G160S
kok 53B G160S
3905 CTG TAT TTG AGG AAC TGC AAG ATC GGC GAA ATC GAG CGG GAG ACC TTC AAG GGA
152 L Y L R N C K I G E I E R E T F K G
LRR3
3959 CTG ACC AAT CTG GTG GAG TTG GAT CTG TCA CAT AAT CTG CTG GTT ACC GTG CCC AGT
170 L T N L V E L D L S H N L L V T V P S
LRR4
4016 TTG GCC CTG GGC CAC ATA CCC TCA CTG CGC GAA CTC ACC CTG GCC TCC AAT CAC ATA
189 L A L G H I P S L R E L T L A S N H I
LRR5
4073 CAC AAA ATC GAG AGC CAG GCC TTC GGG AAC ACA CCA TCG CTG CAC AAA TTG GAT CTG
208 H K I E S Q A F G N T P S L H K L D L
LRR6
4130 TCG CAT TGC GAT ATT CAG ACC ATT TCC GCC CAG GCA TTT GGT GGC CTC CAA GGA TTG
227 S H C D I Q T I S A Q A F G G L Q G L
4187 ACT TTG CTC CGA TTG AAT GGC AAT AAA CTG AGC GAG CTT TTG CCC AAG ACA ATT GAG
246 T L L R L N G N K L S E L L P K T I E
LRR7
C Flank
4244 ACC CTG AGT CGA CTT CAT GGC ATC GAA CTG CAC GAC AAT CCC TGG CTC TGT GAT TGT
265 T L S R L H G I E L H D N P W L C D C
4301 CGA TTG AGG GAC ACG AAG CTC TGG CTG ATG AAG AGG AAC ATA CCC TAT CCG GTG GCT
284 R L R D T K L W L M K R N I P Y P V A
kok 118 P309S
kok 65 P309L
4358 CCG GTT TGC TCG GGT GGC CCC GAA AGG ATT ATC GAT CGC AGC TTT GCG GAT CTG CAT
303 P V C S G G P E R I I D R S F A D L H

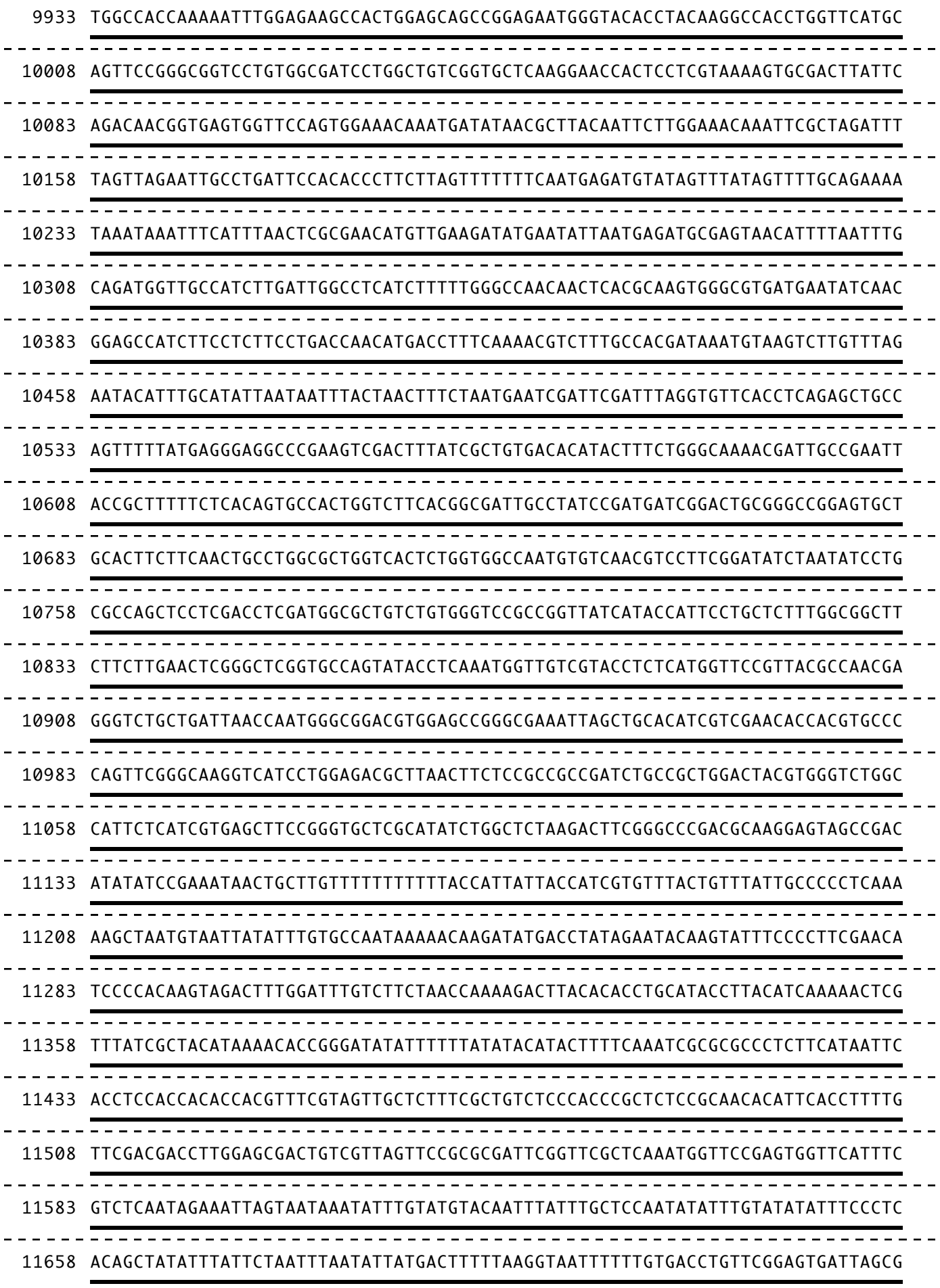

kok 82A P329S
4415 GTG GAT GAG TTT GCC TGC CGA CCG GAG ATG TTG CCC ATA TCG CAT TAT GTG GAG GCG
322 V D E F A C R P E M L P I S H Y V E A
XhoI
kok 176v P35
4472 GCC ATG GGC GAG AAT GCC TCG ATT ACA TGT CGA GCT CGA GCG GTT CCA GCT GCG AAT
341 A M G E N A S I T C R A R A V P A A N
4529 ATC AAC TGG TAC TGG AAC GGA CGG CTG CTG GCC AAC AAT TCC GCC TTC ACC GCG TAC
360 I N W Y W N G R L L A N N S A F T A Y
RNA1 43521 / 4761
W65
4586 CAG AGG ATA CAC ATG TTG GAG CAG GTG GAA GGT GGA TTC GAA AAG CGA TCC AAA CTG
379 Q R I H M L E Q V E G G F E K R S K L
4643 GTG CTG ACC AAC GCA CAG GAA ACG GAT TCC AGT GAG TTC TAC TGC GTG GCC GAG AAT
398 V L T N A Q E T D S S E F Y C V A E N
Tm Swap
4700 CGA GCT GGG ATG GCC GAG GCC AAC TTC ACC CTG CAC GTG AGC ATG AGA GCT GCG GGC
417 R A G M A E A N F T L H V S M R A A G
4757 ATG GCC TCC CTG GGT AGT GGC CAA ATT GTG GGT CTG AGT GCC GCC CTG GTT GCT CTG
436 M A S L G S G Q I V G L S A A L V A L
4814 ATT GTG TTT GCC CTT GGG GTT ATC ATG TGC CTG CTC CTG AGG GTA AAA CGG CAG CCG
455 I V F A L G V I M C L L L R V K R Q P
4871 TAT GTC GAT AGC AAG ACG CCC AAT CAC ATG GAG GTG ATA ACA TCT GTT AAC CAC CAG
474 Y V D S K T P N H M E V I T S V N H Q
4928 AAC TCC ATA ACA AAC AAG ACG CAG CCC GCA ACG GGA AAT GGC AGT ATT GGC GGC GTG
493 N S I T N K T Q P A T G N G S I G G V
4985 GTC ATC GCC AAT GGA GCT GTG GCC AAC ATA ATC GAT GGC GGA GTG GTG CAG GGA GGA
512 V I A N G A V A N I I D G G V V Q G G

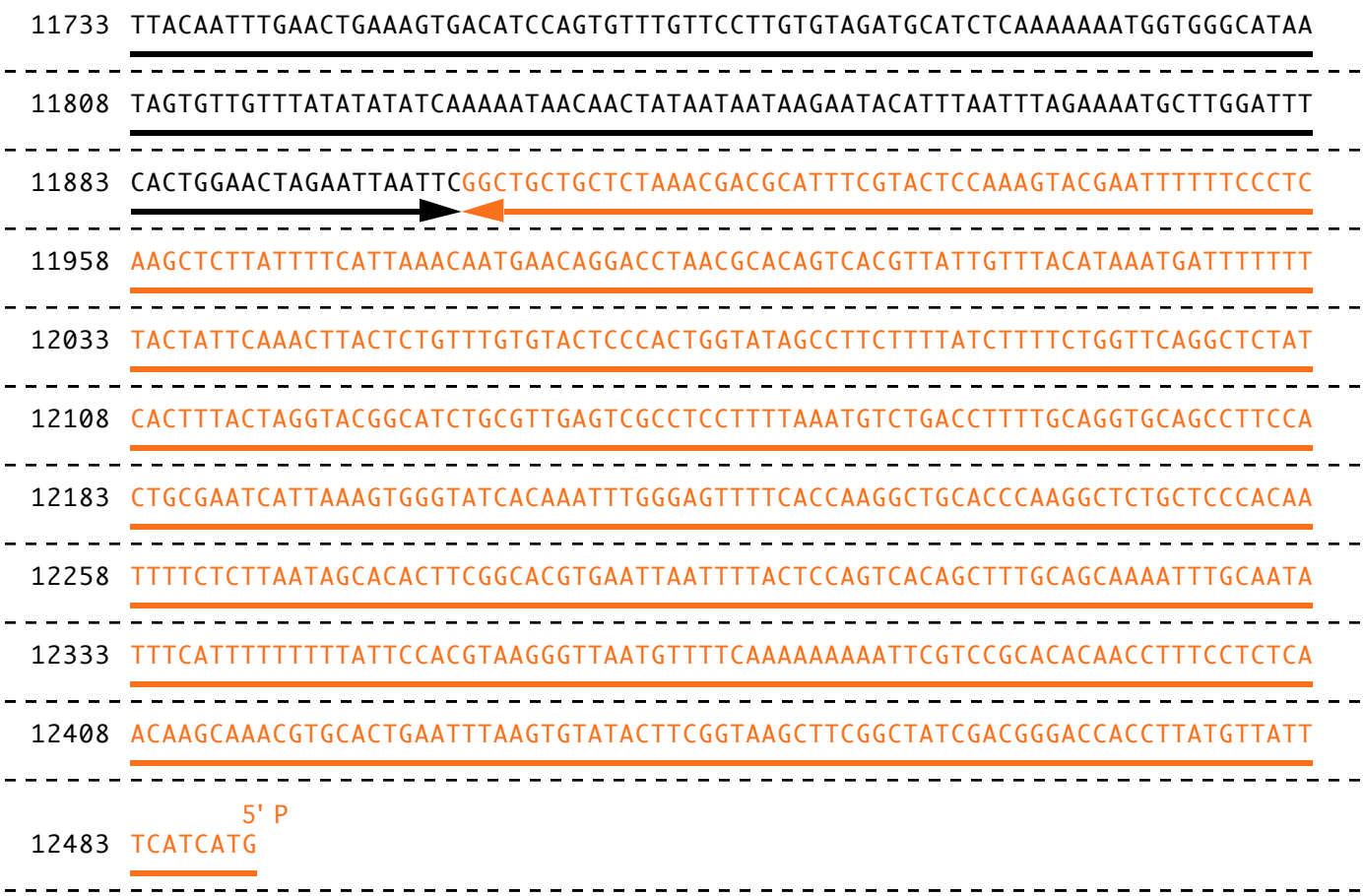

5042 ACT CTG GAG CGG AAA AGC AGC GGA CGG GGA GGT GTA CCG CAT GGA GTT CAC GAT CAG
531 T L E R K S S G R G G V P H G V H D Q
BglII
5099 CGC AGT GCA AAT CCC GTG CAG AAA CCG CCG AGG CTA ACA GAT CTT CCG TAC TCT ACG
550 R S A N P V Q K P P R L T D L P Y S T
5156 CAG GGC TAT GAC AAC AAC GGA AGT GTC CTG TCC ACT GCC TCC TGT TTC ATC TCG CCC
569 Q G Y D N N G S V L S T A S C F I S P
5213 AGT GGA TCC ACC GGA AAC GGT GGC AAC AAT CCT GAT CTC ATT AAT GAT ACC AAA CGT
588 S G S T G N G G N N P D L I N D T K R
5270 TTT GGG AGC GAC GAG TTT GCG GAT CTG AAG ATA CCA CCC ATC AGT GGT GTT GGA GTC
607 F G S D E F A D L K I P P I S G V G V
RNAi-36252
W64
5327 GGC GGC AGT GGG GAG TAT AGT CGC GCC AAC GGC TGC GAT TCC CTG TAT CCT TCG GGT
626 G G S G E Y S R A N G C D S L Y P S G
5384 CTG TGG GAA CAT GGT GCT CCA GTG GGC ACC ACA TCC GCG GAT GAC CTC TTC ATG AAG
645 L W E H G A P V G T T S A D D L F M K
5441 CGC TAC ACC GAC AAG ACG CCC ATC ATA GAC TCC ACA CAG CTG TAC GAC CTT CAT GAG
664 R Y T D K T P I I D S T Q L Y D L H E
BglII
5498 CGA ACG GCG GCC ACG GAT TAT TTT AGC AAG ACA TTC CCG AGA TCT CAC CTC CAG CAG
683 R T A A T D Y F S K T F P R S H L Q Q
5555 GGC ATG ATG ACG GGT GGC GGT GGA GGA ACC TCG ACG GCG TCG ACG GTA ACC ACT AAT
702 G M M T G G G G G T S T A S T V T T N
5612 TTG TCG GGT GGC TCC TCA TCG GGT TAC CCC AAC GAT TAT GGT CTG CCT CTG GTG CCG
721 L S G G S S S G Y P N D Y G L P L V P
5669 GGG GCA GAG CAC CAG CAC AAC CAC CAG CTG CAG ATG CAT CCA CTG CAG CAG CTC CAG
740 G A E H Q H N H Q L Q M H P L Q Q L Q
5726 CAG CAG CTG ACC TCC ACG CTG AAC CAT CAG AAG CAG GAG GGC AGC TCC ACC GGG AGC
759 Q Q L T S T L N H Q K Q E G S S T G S

```
5783 AGT CCG CAC TTC AGT AGC CGC ACA CTG CCA CGC CTG CAC GAG GGC AGT GGC GGG GGC
 778  S   P   H   F   S   S   R   T   L   P   R   L   H   E   G   S   G   G   G
5840 GGC AGT TCA CGG TCG TCG CCG ACG CCA GCG ATT AGC GGT GGC CAT GCC AAC CAG GCG
 797  G   S   S   R   S   S   P   T   P   A   I   S   G   G   H   A   N   Q   A
5897 GCA AAT CCC AGC ACC TCC AGT TCC TCC TGC TCC ATC CTG CCC AAC GGG CAG CCA ATT
 816  A   N   P   S   T   S   S   S   S   C   S   I   L   P   N   G   Q   P   I
5954 AAC GCC AAG ACG ATA CGG GTG TGG CAA AAG GGC GGT GTG CCC GTC CTG CCA CCC GTG
 835  N   A   K   T   I   R   V   W   Q   K   G   G   V   P   V   L   P   P   V
6011 ACG GCG CTG AAA AGG GCC CTG ATC AGC AGC AGC CGG AAT TCG CCG GAC GAG GGA TAC
 854  T   A   L   K   R   A   L   I   S   S   S   R   N   S   P   D   E   G   Y
                                  attB2                                KpnI    EGFP
6068 CAG GAA GGA TGC GGC ACG GAT GTG CAC CCA GCT TTC TTG TAC AAA GTG GTG GTA CCG
 873  Q   E   G   C   G   T   D   V   H   P   A   F   L   Y   K   V   V   V   P
                                          ATG for EGFPN1
6125 CGG GCC CGG GAT CCA CCG GTC GCC ACC ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC
 892  R   A   R   D   P   P   V   A   T   M   V   S   K   G   E   E   L   F   T
6182 GGG GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC
 911  G   V   V   P   I   L   V   E   L   D   G   D   V   N   G   H   K   F   S
6239 GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC
 930  V   S   G   E   G   E   G   D   A   T   Y   G   K   L   T   L   K   F   I
6296 TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC CTG ACC TAC
 949  C   T   T   G   K   L   P   V   P   W   P   T   L   V   T   T   L   T   Y
6353 GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG
 968  G   V   Q   C   F   S   R   Y   P   D   H   M   K   Q   H   D   F   F   K
6410 TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC
 987  S   A   M   P   E   G   Y   V   Q   E   R   T   I   F   F   K   D   D   G
6467 AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC
1006  N   Y   K   T   R   A   E   V   K   F   E   G   D   T   L   V   N   R   I
6524 GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG
1025  E   L   K   G   I   D   F   K   E   D   G   N   I   L   G   H   K   L   E
6581 TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC GGC ATC
1044  Y   N   Y   N   S   H   N   V   Y   I   M   A   D   K   Q   K   N   G   I
6638 AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC
1063  K   V   N   F   K   I   R   H   N   I   E   D   G   S   V   Q   L   A   D
```


6695 CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC
1082 H Y Q Q N T P I G D G P V L L P D N H
6752 TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG
1101 Y L S T Q S A L S K D P N E K R D H M
6809 GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC
1120 V L L E F V T A A G I T L G M D E L Y
SV40 Poly A
XbaI BglII
6866 AAG TAA AGC GGC CGC GAC TCT AGA GATCTTTGTGAAGGAACCTTACTTCTGTGGTGTGACATAATTG
1139 K • S G R D S R
6933 GACAAACTACCTACAGAGATTTAAAGCTCTAAGGTAAATATAAAATTTTTAAGTGTATAATGTGTTAAACTACTG
7008 ATTCTAATTGTTTGTGTATTTTAGATTCCAACCTATGGAACTGATGAATGGGAGCAGTGGTGGAATGCCTTTAAT
7083 GAGGAAAACCTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGACTCTCAACATTCTACT
7158 CCTCCAAAAAAGAAGAGAAAGGTAGAAGACCCCAAGGACTTTCCTTCAGAATTGCTAAGTTTTTTGAGTCATGCT
7233 GTGTTTAGTAATAGAACTCTTGCTTGCTTTGCTATTTACACCACAAAGGAAAAAGCTGCACTGCTATACAAGAAA
7308 ATTATGGAAAAATATTCTGTAACCTTTATAAGTAGGCATAACAGTTATAATCATAACATACTGTTTTTTCTTACT
7383 CCACACAGGCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAATTGTGTACCTTTAGCTTTTTAATTTGTAAA
7458 GGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAGATCATAATCAGCCATACCACATTTGTAGAGGT
7533 TTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAA
HpaI
7608 CTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTC
white gene
BamHI
7683 ACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGGATCGGATCCACTAGAAGGCC
7758 TTAGTATGTATGTAAGTTAATAAAACCCTTTTTTGGAGAATGTAGATTTAAAAAAACATATTTTTTTTTTATTTT
7833 TTACTGCACTGGACATCATTGAACTTATCTGATCAGTTTTAAATTTACTTCGATCCAAGGGTATTTGAAGTACCA
7908 GGTTCTTTCGATTACCTCTCACTCAAAATGACATTCCACTCAAAGTCAGCGCTGTTTGCCTCCTTCTCTGTCCAC
7983 AGAAATATCGCCGTCTCTTTCGCCGCTGCGTCCGCTATCTCTTTCGCCACCGTTTGTAGCGTTACCTAGCGTCAA
8058 TGTCCGCCTTCAGTTGCACTTTGTCAGCGGTTTCGTGACGAAGCTCCAAGCGGTTTACGCCATCAATTAAACACA

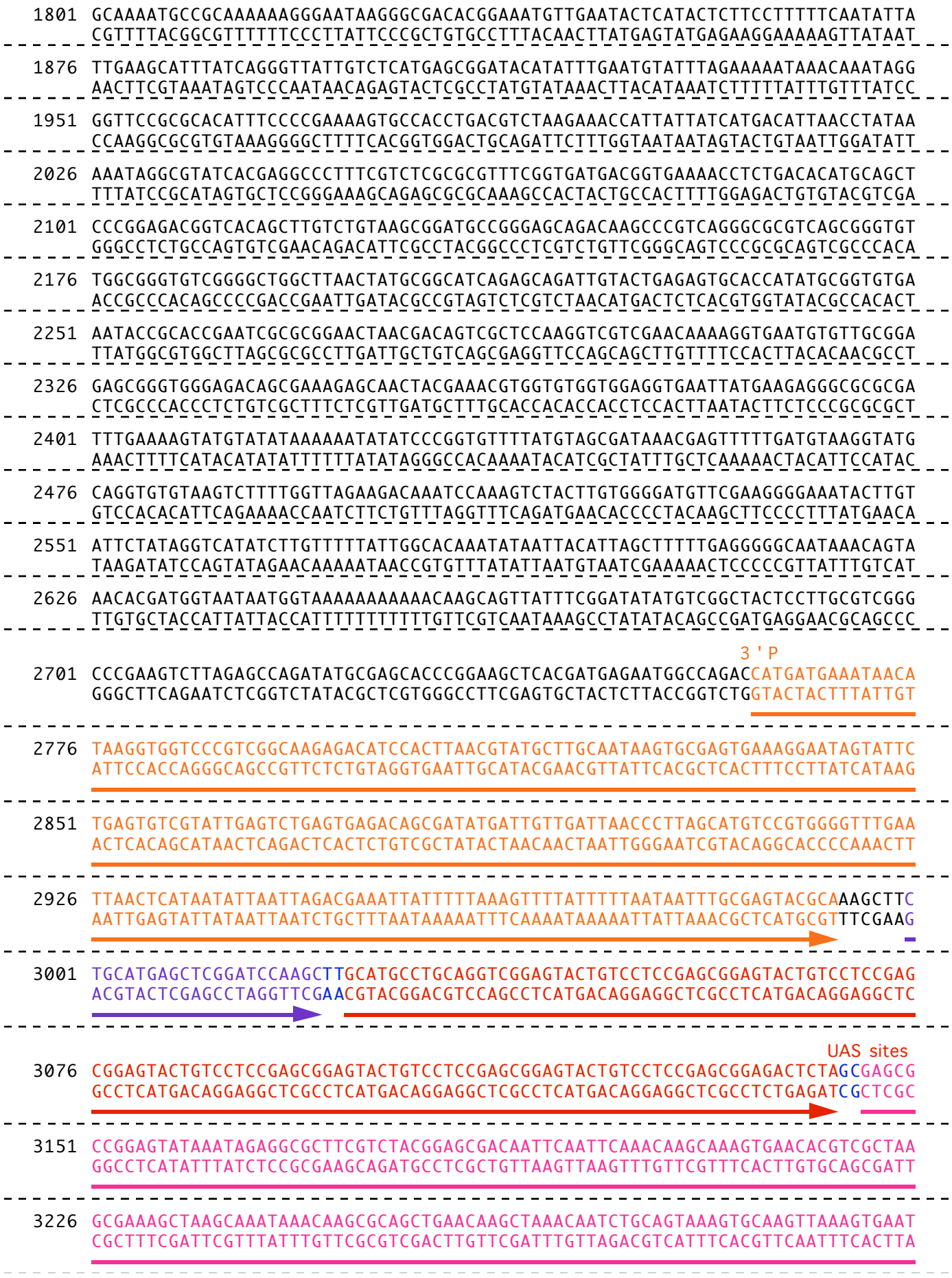

8133 AAGTGCTGTGCCAAAACTCCTCTCGCTTCTTATTTTTGTTTGTTTTTTGAGTGATTGGGGTGGTGATTGGTTTTG
8208 GGTGGGTAAGCAGGGGAAAGTGTGAAAAATCCCGGCAATGGGCCAAGAGGATCAGGAGCTATTAATTCGCGGAGG
8283 CAGCAAACACCCATCTGCCGAGCATCTGAACAATGTGAGTAGTACATGTGCATACATCTTAAGTTCACTTGATCT
8358 ATAGGAACTGCGATTGCAACATCAAATTGTCTGCGGCGTGAGAACTGCGACCCACAAAAATCCCAAACCGCAATC
8433 GCACAAACAAATAGTGACACGAAACAGATTATTCTGGTAGCTGTGCTCGCTATATAAGACAATTTTTAAGATCAT
8508 ATCATGATCAAGACATCTAAAGGCATTCATTTTCGACTACATTCTTTTTTACAAAAAATATAACAACCAGATATT
8583 TTAAGCTGATCCTAGATGCACAAAAAATAAATAAAAGTATAAACCTACTTCGTAGGATACTTCGTTTTGTTCGGG
8658 GTTAGATGAGCATAACGCTTGTAGTTGATATTTGAGATCCCCTATCATTGCAGGGTGACAGCGGAGCGGCTTCGC
8733 AGAGCTGCATTAACCAGGGCTTCGGGCAGGCCAAAAACTACGGCACGCTCCTGCCACCCAGTCCGCCGGAGGACT
8808 CCGGTTCAGGGAGCGGCCAACTAGCCGAGAACCTCACCTATGCCTGGCACAATATGGACATCTTTGGGGCGGTCA
8883 ATCAGCCGGGCTCCGGATGGCGGCAGCTGGTCAACCGGACACGCGGACTATTCTGCAACGAGCGACACATACCGG
8958 CGCCCAGGAAACATTTGCTCAAGAACGGTGAGTTTCTATTCGCAGTCGGCTGATCTGTGTGAAATCTTAATAAAG
9033 GGTCCAATTACCAATTTGAAACTCAGTTTGCGGCGTGGCCTATCCGGGCGAACTTTTGGCCGTGATGGGCAGTTC
9108 CGGTGCCGGAAAGACGACCCTGCTGAATGCCCTTGCCTTTCGATCGCCGCAGGGCATCCAAGTATCGCCATCCGG
9183 GATGCGACTGCTCAATGGCCAACCTGTGGACGCCAAGGAGATGCAGGCCAGGTGCGCCTATGTCCAGCAGGATGA
9258 CCTCTTTATCGGCTCCCTAACGGCCAGGGAACACCTGATTTTCCAGGCCATGGTGCGGATGCCACGACATCTGAC
9333 CTATCGGCAGCGAGTGGCCCGCGTGGATCAGGTGATCCAGGAGCTTTCGCTCAGCAAATGTCAGCACACGATCAT
9408 CGGTGTGCCCGGCAGGGTGAAAGGTCTGTCCGGCGGAGAAAGGAAGCGTCTGGCATTCGCCTCCGAGGCACTAAC
9483 CGATCCGCCGCTTCTGATCTGCGATGAGCCCACCTCCGGACTGGACTCATTTACCGCCCACAGCGTCGTCCAGGT
9558 GCTGAAGAAGCTGTCGCAGAAGGGCAAGACCGTCATCCTGACCATTCATCAGCCGTCTTCCGAGCTGTTTGAGCT
9633 CTTTGACAAGATCCTTCTGATGGCCGAGGGCAGGGTAGCTTTCTTGGGCACTCCCAGCGAAGCCGTCGACTTCTT
9708 TTCCTAGTGAGTTCGATGTGTTTATTAAGGGTATCTAGCATTACATTACATCTCAACTCCTATCCAGCGTGGGTG
9783 CCCAGTGTCCTACCAACTACAATCCGGCGGACTTTTACGTACAGGTGTTGGCCGTTGTGCCCGGACGGGAGATCG
9858 AGTCCCGTGATCGGATCGCCAAGATATGCGACAATTTTGCTATTAGCAAAGTAGCCCGGGATATGGAGCAGTTGT

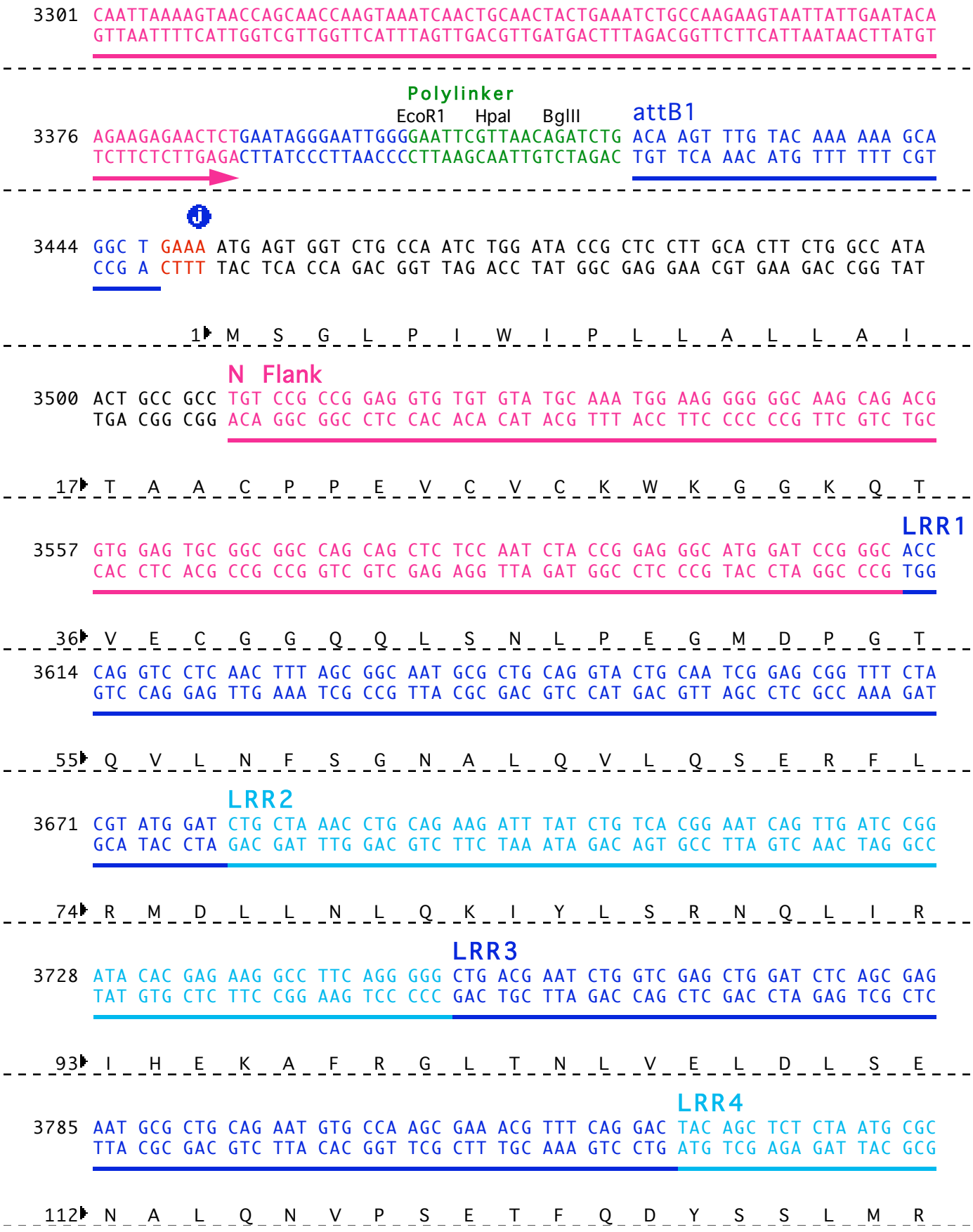

9933 TGGCCACCAAAAATTTGGAGAAGCCACTGGAGCAGCCGGAGAATGGGTACACCTACAAGGCCACCTGGTTCATGC
10008 AGTTCCGGGCGGTCCTGTGGCGATCCTGGCTGTCGGTGCTCAAGGAACCACTCCTCGTAAAAGTGCGACTTATTC
10083 AGACAACGGTGAGTGGTTCCAGTGGAAACAAATGATATAACGCTTACAATTCTTGGAAACAAATTCGCTAGATTT
10158 TAGTTAGAATTGCCTGATTCCACACCCTTCTTAGTTTTTTTCAATGAGATGTATAGTTTATAGTTTTGCAGAAAA
10233 TAAATAAATTTCATTTAACTCGCGAACATGTTGAAGATATGAATATTAATGAGATGCGAGTAACATTTTAATTTG
10308 CAGATGGTTGCCATCTTGATTGGCCTCATCTTTTTGGGCCAACAACTCACGCAAGTGGGCGTGATGAATATCAAC
10383 GGAGCCATCTTCCTCTTCCTGACCAACATGACCTTTCAAAACGTCTTTGCCACGATAAATGTAAGTCTTGTTTAG
10458 AATACATTTGCATATTAATAATTTACTAACTTTCTAATGAATCGATTCGATTTAGGTGTTCACCTCAGAGCTGCC
10533 AGTTTTTATGAGGGAGGCCCGAAGTCGACTTTATCGCTGTGACACATACTTTCTGGGCAAAACGATTGCCGAATT
10608 ACCGCTTTTTCTCACAGTGCCACTGGTCTTCACGGCGATTGCCTATCCGATGATCGGACTGCGGGCCGGAGTGCT
10683 GCACTTCTTCAACTGCCTGGCGCTGGTCACTCTGGTGGCCAATGTGTCAACGTCCTTCGGATATCTAATATCCTG
10758 CGCCAGCTCCTCGACCTCGATGGCGCTGTCTGTGGGTCCGCCGGTTATCATACCATTCCTGCTCTTTGGCGGCTT
10833 CTTCTTGAACTCGGGCTCGGTGCCAGTATACCTCAAATGGTTGTCGTACCTCTCATGGTTCCGTTACGCCAACGA
10908 GGGTCTGCTGATTAACCAATGGGCGGACGTGGAGCCGGGCGAAATTAGCTGCACATCGTCGAACACCACGTGCCC
10983 CAGTTCGGGCAAGGTCATCCTGGAGACGCTTAACTTCTCCGCCGCCGATCTGCCGCTGGACTACGTGGGTCTGGC
11058 CATTCTCATCGTGAGCTTCCGGGTGCTCGCATATCTGGCTCTAAGACTTCGGGCCCGACGCAAGGAGTAGCCGAC
11133 ATATATCCGAAATAACTGCTTGTTTTTTTTTTTACCATTATTACCATCGTGTTTACTGTTTATTGCCCCCTCAAA
11208 AAGCTAATGTAATTATATTTGTGCCAATAAAAACAAGATATGACCTATAGAATACAAGTATTTCCCCTTCGAACA
11283 TCCCCACAAGTAGACTTTGGATTTGTCTTCTAACCAAAAGACTTACACACCTGCATACCTTACATCAAAAACTCG
11358 TTTATCGCTACATAAAACACCGGGATATATTTTTTATATACATACTTTTCAAATCGCGCGCCCTCTTCATAATTC
11433 ACCTCCACCACACCACGTTTCGTAGTTGCTCTTTCGCTGTCTCCCACCCGCTCTCCGCAACACATTCACCTTTTG
11508 TTCGACGACCTTGGAGCGACTGTCGTTAGTTCCGCGCGATTCGGTTCGCTCAAATGGTTCCGAGTGGTTCATTTC
11583 GTCTCAATAGAAATTAGTAATAAATATTTGTATGTACAATTTATTTGCTCCAATATATTTGTATATATTTCCCTC
11658 ACAGCTATATTTATTCTAATTTAATATTATGACTTTTTAAGGTAATTTTTTGTGACCTGTTCGGAGTGATTAGCG

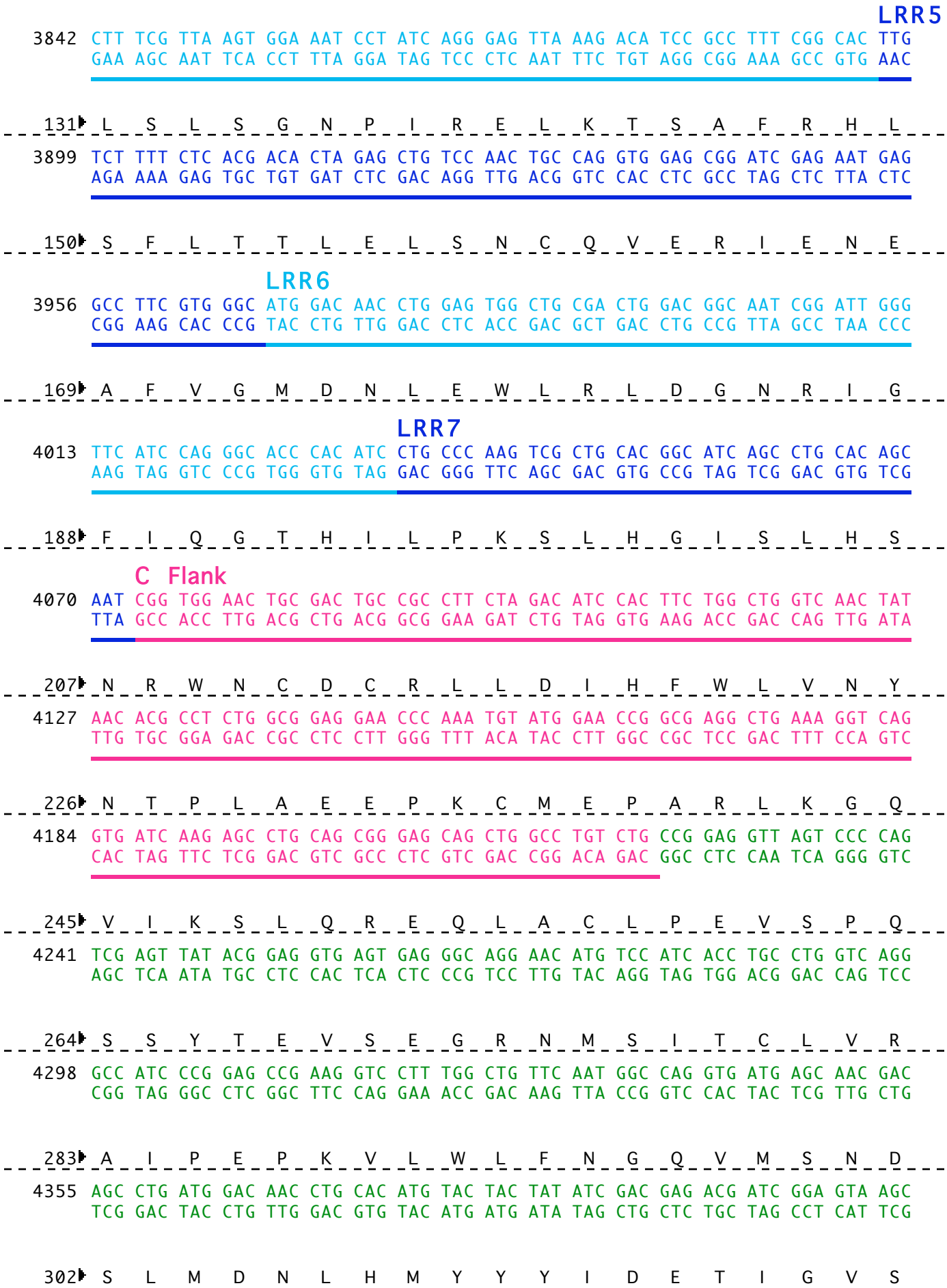

11733 TTACAATTTGAACTGAAAGTGACATCCAGTGTTTGTTCCTTGTGTAGATGCATCTCAAAAAAATGGTGGGCATAA
11808 TAGTGTTGTTTATATATATCAAAAATAACAACTATAATAATAAGAATACATTTAATTTAGAAAATGCTTGGATTT
11883 CACTGGAACTAGAATTAATTCGGCTGCTGCTCTAAACGACGCATTTCGTACTCCAAAGTACGAATTTTTTCCCTC
11958 AAGCTCTTATTTTCATTAAACAATGAACAGGACCTAACGCACAGTCACGTTATTGTTTACATAAATGATTTTTTT
12033 TACTATTCAAACTTACTCTGTTTGTGTACTCCCACTGGTATAGCCTTCTTTTATCTTTTCTGGTTCAGGCTCTAT
12108 CACTTTACTAGGTACGGCATCTGCGTTGAGTCGCCTCCTTTTAAATGTCTGACCTTTTGCAGGTGCAGCCTTCCA
12183 CTGCGAATCATTAAAGTGGGTATCACAAATTTGGGAGTTTTCACCAAGGCTGCACCCAAGGCTCTGCTCCCACAA
12258 TTTTCTCTTAATAGCACACTTCGGCACGTGAATTAATTTTACTCCAGTCACAGCTTTGCAGCAAAATTTGCAATA
12333 TTTCATTTTTTTTTATTCCACGTAAGGGTTAATGTTTTCAAAAAAAAATTCGTCCGCACACAACCTTTCCTCTCA
12408 ACAAGCAAACGTGCACTGAATTTAAGTGTATACTTCGGTAAGCTTCGGCTATCGACGGGACCACCTTATGTTATT
5' P
12483 TCATCATG

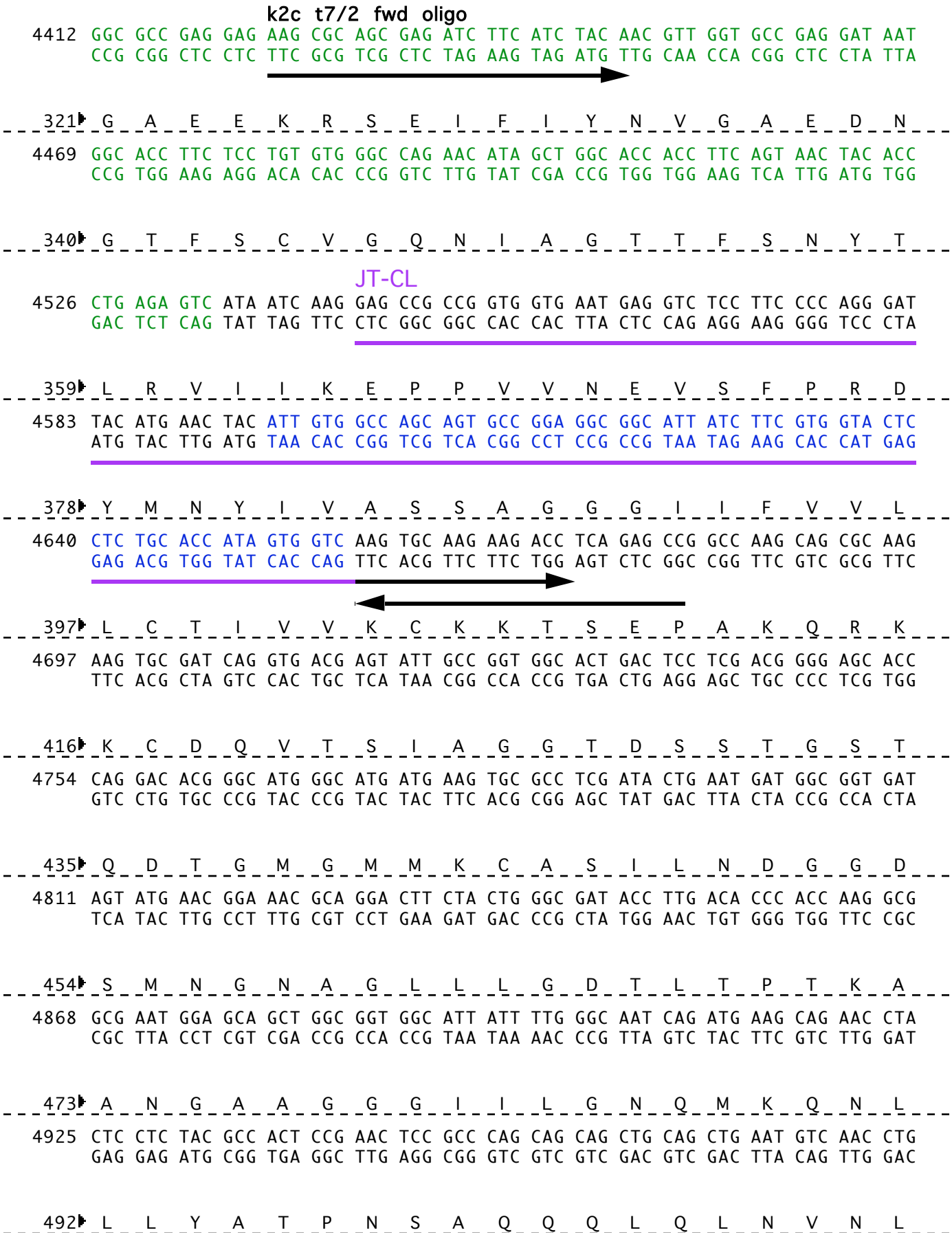
UAS Kek2 GFP

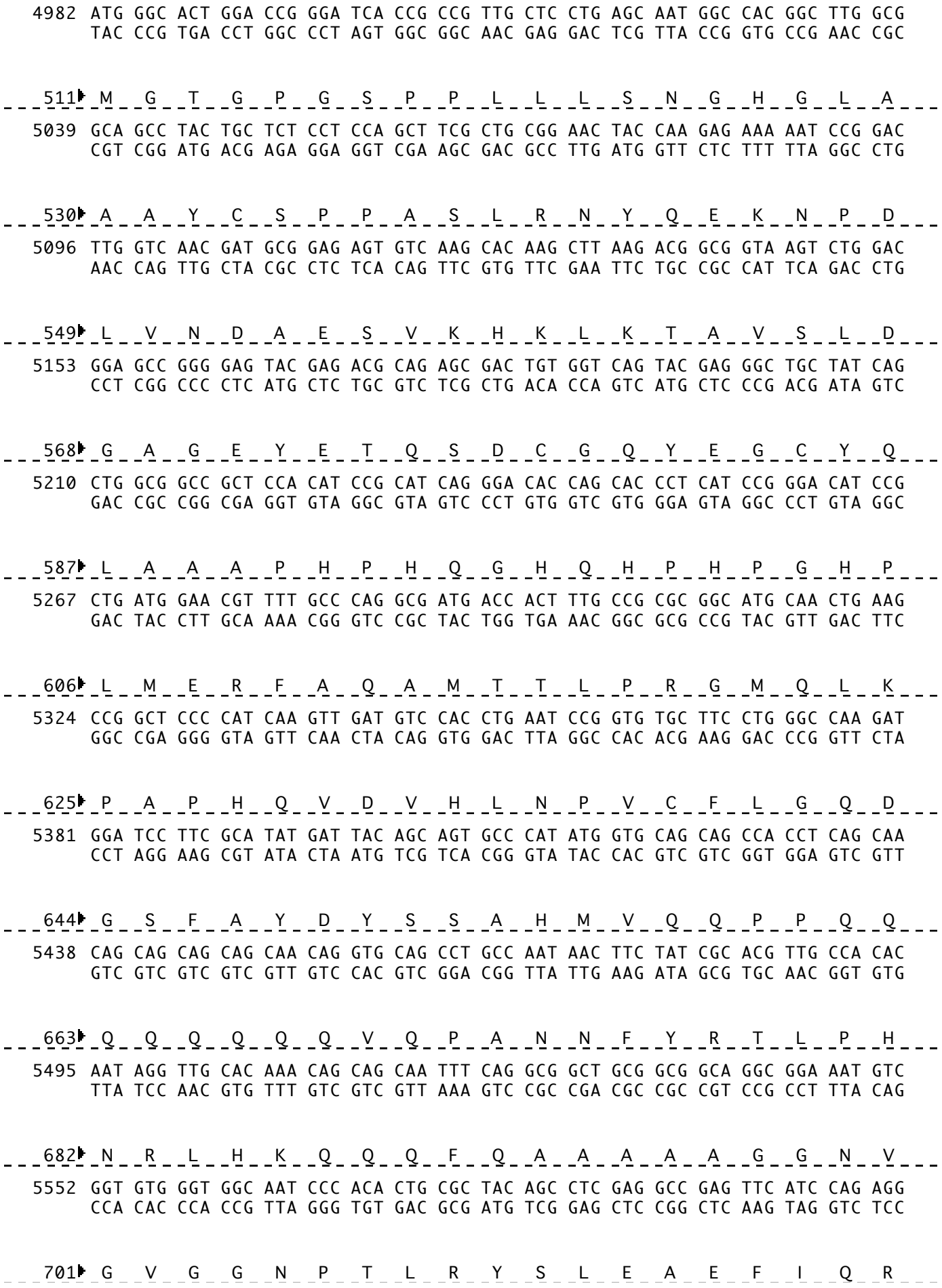
```
     1 GGCCAGACCCACGTAGTCCAGCGGCAGATCGGCGGCGGAGAAGTTAAGCGTCTCCAGGATGACCTTGCCCGAACT
       CCGGTCTGGGTGCATCAGGTCGCCGTCTAGCCGCCGCCTCTTCAATTCGCAGAGGTCCTACTGGAACGGGCTTGA

    76 GGGGCACGTGGTGTTCGACGATGTGCAGCTAATTTCGCCCGGCTCCACGTCCGCCCATTGGTTAATCAGCAGACC
       CCCCGTGCACCACAAGCTGCTACACGTCGATTAAAGCGGGCCGAGGTGCAGGCGGGTAACCAATTAGTCGTCTGG

   151 CTCGTTGGCGTAACGGAACCATGAGAGGTACGACAACCATTTGAGGTATACTGGCACCGAGCCCGAGTTCAAGAA
       GAGCAACCGCATTGCCTTGGTACTCTCCATGCTGTTGGTAAACTCCATATGACCGTGGCTCGGGCTCAAGTTCTT

   226 GAAGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA
       CTTCCGCAAAAAGGTATCCGAGGCGGGGGGACTGCTCGTAGTGTTTTTAGCTGCGAGTTCAGTCTCCACCGCTTT

   301 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC
       GGGCTGTCCTGATATTTCTATGGTCCGCAAAGGGGGACCTTCGAGGGAGCACGCGAGAGGACAAGGCTGGGACGG

   376 GCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCT
       CGAATGGCCTATGGACAGGCGGAAAGAGGGAAGCCCTTCGCACCGCGAAAGAGTTACGAGTGCGACATCCATAGA

   451 CAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTT
       GTCAAGCCACATCCAGCAAGCGAGGTTCGACCCGACACACGTGCTTGGGGGGCAAGTCGGGCTGGCGACGCGGAA

   526 ATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG
       TAGGCCATTGATAGCAGAACTCAGGTTGGGCCATTCTGTGCTGAATAGCGGTGACCGTCGTCGGTGACCATTGTC

   601 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAG
       CTAATCGTCTCGCTCCATACATCCGCCACGATGTCTCAAGAACTTCACCACCGGATTGATGCCGATGTGATCTTC

   676 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA
       CTGTCATAAACCATAGACGCGAGACGACTTCGGTCAATGGAAGCCTTTTTCTCAACCATCGAGAACTAGGCCGTT

   751 ACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGA
       TGTTTGGTGGCGACCATCGCCACCAAAAAAACAAACGTTCGTCGTCTAATGCGCGTCTTTTTTTCCTAGAGTTCT

   826 AGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG
       TCTAGGAAACTAGAAAAGATGCCCCAGACTGCGAGTCACCTTGCTTTTGAGTGCAATTCCCTAAAACCAGTACTC

   901 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGA
       TAATAGTTTTTCCTAGAAGTGGATCTAGGAAAATTTAATTTTTACTTCAAAATTTAGTTAGATTTCATATATACT

   976 GTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATC
       CATTTGAACCAGACTGTCAATGGTTACGAATTAGTCACTCCGTGGATAGAGTCGCTAGACAGATAAAGCAAGTAG

  1051 CATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAAT
       GTATCAACGGACTGAGGGGCAGCACATCTATTGATGCTATGCCCTCCCGAATGGTAGACCGGGGTCACGACGTTA

  1126 GATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG
       CTATGGCGCTCTGGGTGCGAGTGGCCGAGGTCTAAATAGTCGTTATTTGGTCGGTCGGCCTTCCCGGCTCGCGTC

  1201 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCC
       TTCACCAGGACGTTGAAATAGGCGGAGGTAGGTCAGATAATTAACAACGGCCCTTCGATCTCATTCATCAAGCGG

  1276 AGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTC
       TCAATTATCAAACGCGTTGCAACAACGGTAACGATGTCCGTAGCACCACAGTGCGAGCAGCAAACCATACCGAAG

  1351 ATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTT
       TAAGTCGAGGCCAAGGGTTGCTAGTTCCGCTCAATGTACTAGGGGGTACAACACGTTTTTTCGCCAATCGAGGAA

  1426 CGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC
       GCCAGGAGGCTAGCAACAGTCTTCATTCAACCGGCGTCACAATAGTGAGTACCAATACCGTCGTGACGTATTAAG

  1501 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG
       AGAATGACAGTACGGTAGGCATTCTACGAAAAGACACTGACCACTCATGAGTTGGTTCAGTAAGACTCTTATCAC

  1576 TATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGT
       ATACGCCGCTGGCTCAACGAGAACGGGCCGCAGTTGTGCCCTATTATGGCGCGGTGTATCGTCTTGAAATTTTCA

  1651 GCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTA
       CGAGTAGTAACCTTTTGCAAGAAGCCCCGCTTTTGAGAGTTCCTAGAATGGCGACAACTCTAGGTCAAGCTACAT

  1726 ACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG
       TGGGTGAGCACGTGGGTTGACTAGAAGTCGTAGAAAATGAAAGTGGTCGCAAAGACCCACTCGTTTTTGTCCTTC
```

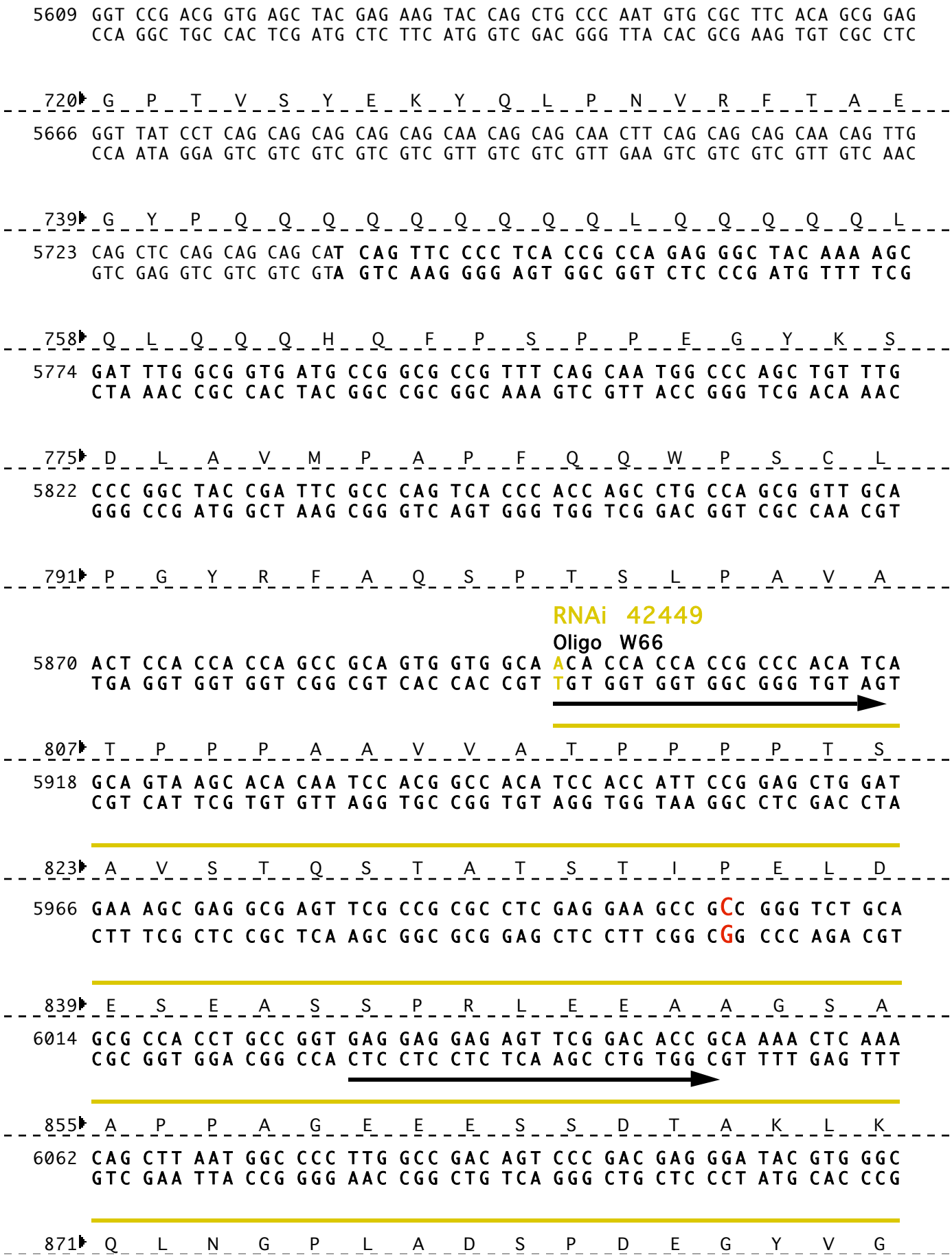

1801 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTA
CGTTTTACGGCGTTTTTTCCCTTATTCCCGCTGTGCCTTTACAACTTATGAGTATGAGAAGGAAAAAGTTATAAT
1876 TTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGG
AACTTCGTAAATAGTCCCAATAACAGAGTACTCGCCTATGTATAAACTTACATAAATCTTTTTATTTGTTTATCC
1951 GGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAA
CCAAGGCGCGTGTAAAGGGGCTTTTCACGGTGGACTGCAGATTCTTTGGTAATAATAGTACTGTAATTGGATATT
2026 AAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCT
TTTATCCGCATAGTGCTCCGGGAAAGCAGAGCGCGCAAAGCCACTACTGCCACTTTTGGAGACTGTGTACGTCGA
2101 CCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGT
GGGCCTCTGCCAGTGTCGAACAGACATTCGCCTACGGCCCTCGTCTGTTCGGGCAGTCCCGCGCAGTCGCCCACA
2176 TGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGA
ACCGCCCACAGCCCCGACCGAATTGATACGCCGTAGTCTCGTCTAACATGACTCTCACGTGGTATACGCCACACT
2251 AATACCGCACCGAATCGCGCGGAACTAACGACAGTCGCTCCAAGGTCGTCGAACAAAAGGTGAATGTGTTGCGGA
TTATGGCGTGGCTTAGCGCGCCTTGATTGCTGTCAGCGAGGTTCCAGCAGCTTGTTTTCCACTTACACAACGCCT
2326 GAGCGGGTGGGAGACAGCGAAAGAGCAACTACGAAACGTGGTGTGGTGGAGGTGAATTATGAAGAGGGCGCGCGA
CTCGCCCACCCTCTGTCGCTTTCTCGTTGATGCTTTGCACCACACCACCTCCACTTAATACTTCTCCCGCGCGCT
2401 TTTGAAAAGTATGTATATAAAAAATATATCCCGGTGTTTTATGTAGCGATAAACGAGTTTTTGATGTAAGGTATG
AAACTTTTCATACATATATTTTTTATATAGGGCCACAAAATACATCGCTATTTGCTCAAAAACTACATTCCATAC
2476 CAGGTGTGTAAGTCTTTTGGTTAGAAGACAAATCCAAAGTCTACTTGTGGGGATGTTCGAAGGGGAAATACTTGT
GTCCACACATTCAGAAAACCAATCTTCTGTTTAGGTTTCAGATGAACACCCCTACAAGCTTCCCCTTTATGAACA
2551 ATTCTATAGGTCATATCTTGTTTTTATTGGCACAAATATAATTACATTAGCTTTTTGAGGGGGCAATAAACAGTA
TAAGATATCCAGTATAGAACAAAAATAACCGTGTTTATATTAATGTAATCGAAAAACTCCCCCGTTATTTGTCAT
2626 AACACGATGGTAATAATGGTAAAAAAAAAAACAAGCAGTTATTTCGGATATATGTCGGCTACTCCTTGCGTCGGG
TTGTGCTACCATTATTACCATTTTTTTTTTTGTTCGTCAATAAAGCCTATATACAGCCGATGAGGAACGCAGCCC
3 ' P
2701 CCCGAAGTCTTAGAGCCAGATATGCGAGCACCCGGAAGCTCACGATGAGAATGGCCAGACCATGATGAAATAACA
GGGCTTCAGAATCTCGGTCTATACGCTCGTGGGCCTTCGAGTGCTACTCTTACCGGTCTGGTACTACTTTATTGT
2776 TAAGGTGGTCCCGTCGGCAAGAGACATCCACTTAACGTATGCTTGCAATAAGTGCGAGTGAAAGGAATAGTATTC
ATTCCACCAGGGCAGCCGTTCTCTGTAGGTGAATTGCATACGAACGTTATTCACGCTCACTTTCCTTATCATAAG
2851 TGAGTGTCGTATTGAGTCTGAGTGAGACAGCGATATGATTGTTGATTAACCCTTAGCATGTCCGTGGGGTTTGAA
ACTCACAGCATAACTCAGACTCACTCTGTCGCTATACTAACAACTAATTGGGAATCGTACAGGCACCCCAAACTT
2926 TTAACTCATAATATTAATTAGACGAAATTATTTTTAAAGTTTTATTTTTAATAATTTGCGAGTACGCAAAGCTTC
AATTGAGTATTATAATTAATCTGCTTTAATAAAAATTTCAAAATAAAAATTATTAAACGCTCATGCGTTTCGAAG
3001 TGCATGAGCTCGGATCCAAGCTTGCATGCCTGCAGGTCGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAG
ACGTACTCGAGCCTAGGTTCGAACGTACGGACGTCCAGCCTCATGACAGGAGGCTCGCCTCATGACAGGAGGCTC
UAS sites
3076 CGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAGCGGAGACTCTAGCGAGCG
GCCTCATGACAGGAGGCTCGCCTCATGACAGGAGGCTCGCCTCATGACAGGAGGCTCGCCTCTGAGATCGCTCGC
3151 CCGGAGTATAAATAGAGGCGCTTCGTCTACGGAGCGACAATTCAATTCAAACAAGCAAAGTGAACACGTCGCTAA
GGCCTCATATTTATCTCCGCGAAGCAGATGCCTCGCTGTTAAGTTAAGTTTGTTCGTTTCACTTGTGCAGCGATT
3226 GCGAAAGCTAAGCAAATAAACAAGCGCAGCTGAACAAGCTAAACAATCTGCAGTAAAGTGCAAGTTAAAGTGAAT
CGCTTTCGATTCGTTTATTTGTTCGCGTCGACTTGTTCGATTTGTTAGACGTCATTTCACGTTCAATTTCACTTA

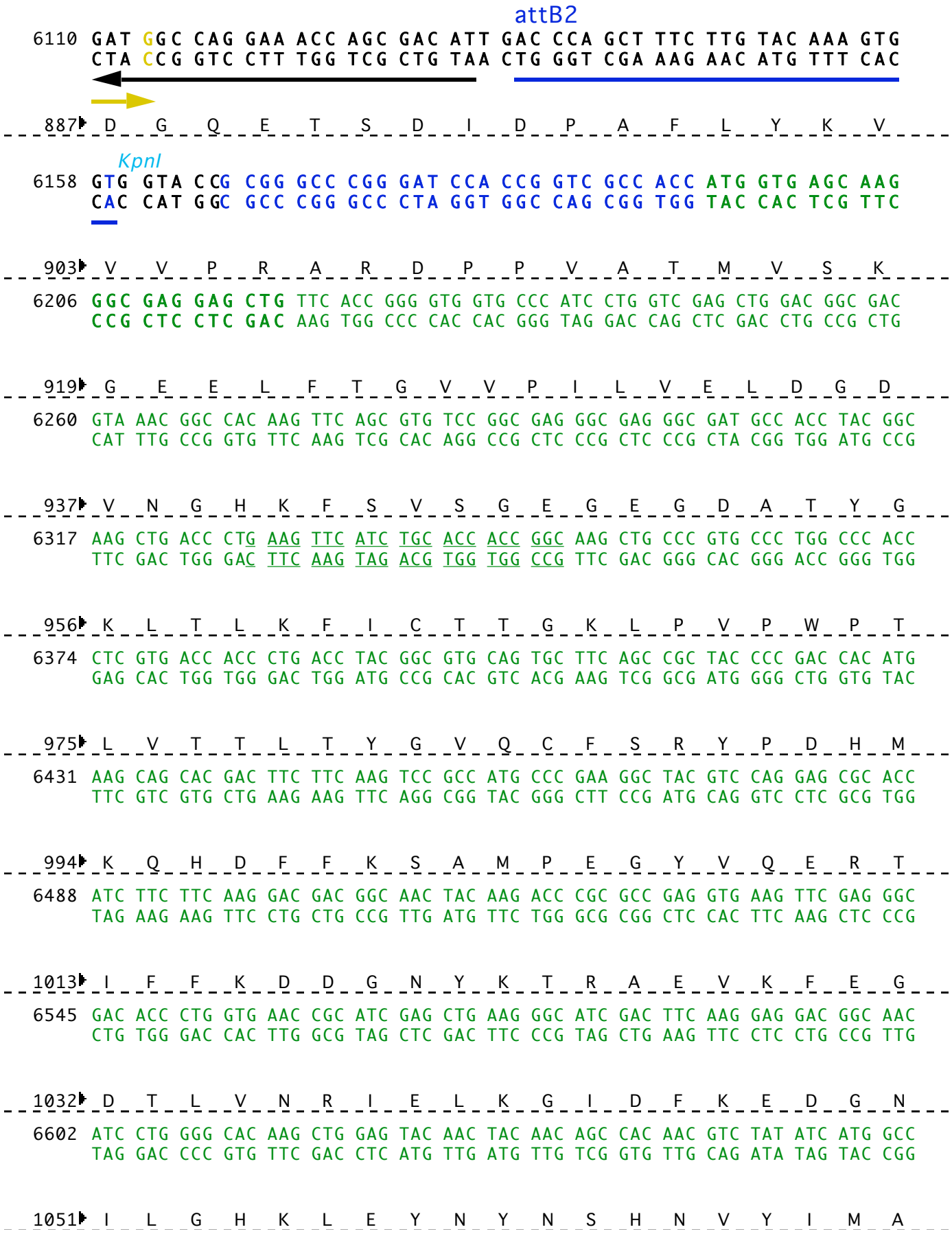

3301 CAATTAAAAGTAACCAGCAACCAAGTAAATCAACTGCAACTACTGAAATCTGCCAAGAAGTAATTATTGAATACA
GTTAATTTTCATTGGTCGTTGGTTCATTTAGTTGACGTTGATGACTTTAGACGGTTCTTCATTAATAACTTATGT
Polylinker
EcoR1 HpaI BglII
attB1
3376 AGAAGAGAACTCTGAATAGGGAATTGGGGAATTCGTTAACAGATCTG ACA AGT TTG TAC AAA AAA GCA
TCTTCTCTTGAGACTTATCCCTTAACCCCTTAAGCAATTGTCTAGAC TGT TCA AAC ATG TTT TTT CGT
3444 GGC T GAAA ATG AGT GGT CTG CCA ATC TGG ATA CCG CTC CTT GCA CTT CTG GCC ATA
CCG A CTTT TAC TCA CCA GAC GGT TAG ACC TAT GGC GAG GAA CGT GAA GAC CGG TAT
1 M S G L P I W I P L L A L L A I
N Flank
3500 ACT GCC GCC TGT CCG CCG GAG GTG TGT GTA TGC AAA TGG AAG GGG GGC AAG CAG ACG
TGA CGG CGG ACA GGC GGC CTC CAC ACA CAT ACG TTT ACC TTC CCC CCG TTC GTC TGC
17 T A A C P P E V C V C K W K G G K Q T
LRR1
3557 GTG GAG TGC GGC GGC CAG CAG CTC TCC AAT CTA CCG GAG GGC ATG GAT CCG GGC ACC
CAC CTC ACG CCG CCG GTC GTC GAG AGG TTA GAT GGC CTC CCG TAC CTA GGC CCG TGG
36 V E C G G Q Q L S N L P E G M D P G T
3614 CAG GTC CTC AAC TTT AGC GGC AAT GCG CTG CAG GTA CTG CAA TCG GAG CGG TTT CTA
GTC CAG GAG TTG AAA TCG CCG TTA CGC GAC GTC CAT GAC GTT AGC CTC GCC AAA GAT
55 Q V L N F S G N A L Q V L Q S E R F L
LRR2
3671 CGT ATG GAT CTG CTA AAC CTG CAG AAG ATT TAT CTG TCA CGG AAT CAG TTG ATC CGG
GCA TAC CTA GAC GAT TTG GAC GTC TTC TAA ATA GAC AGT GCC TTA GTC AAC TAG GCC
74 R M D L L N L Q K I Y L S R N Q L I R
LRR3
3728 ATA CAC GAG AAG GCC TTC AGG GGG CTG ACG AAT CTG GTC GAG CTG GAT CTC AGC GAG
TAT GTG CTC TTC CGG AAG TCC CCC GAC TGC TTA GAC CAG CTC GAC CTA GAG TCG CTC
93 I H E K A F R G L T N L V E L D L S E
LRR4
3785 AAT GCG CTG CAG AAT GTG CCA AGC GAA ACG TTT CAG GAC TAC AGC TCT CTA ATG CGC
TTA CGC GAC GTC TTA CAC GGT TCG CTT TGC AAA GTC CTG ATG TCG AGA GAT TAC GCG
112 N A L Q N V P S E T F Q D Y S S L M R


LRR5
3842 CTT TCG TTA AGT GGA AAT CCT ATC AGG GAG TTA AAG ACA TCC GCC TTT CGG CAC TTG
GAA AGC AAT TCA CCT TTA GGA TAG TCC CTC AAT TTC TGT AGG CGG AAA GCC GTG AAC
131 L S L S G N P I R E L K T S A F R H L
3899 TCT TTT CTC ACG ACA CTA GAG CTG TCC AAC TGC CAG GTG GAG CGG ATC GAG AAT GAG
AGA AAA GAG TGC TGT GAT CTC GAC AGG TTG ACG GTC CAC CTC GCC TAG CTC TTA CTC
150 S F L T T L E L S N C Q V E R I E N E
LRR6
3956 GCC TTC GTG GGC ATG GAC AAC CTG GAG TGG CTG CGA CTG GAC GGC AAT CGG ATT GGG
CGG AAG CAC CCG TAC CTG TTG GAC CTC ACC GAC GCT GAC CTG CCG TTA GCC TAA CCC
169 A F V G M D N L E W L R L D G N R I G
LRR7
4013 TTC ATC CAG GGC ACC CAC ATC CTG CCC AAG TCG CTG CAC GGC ATC AGC CTG CAC AGC
AAG TAG GTC CCG TGG GTG TAG GAC GGG TTC AGC GAC GTG CCG TAG TCG GAC GTG TCG
188 F I Q G T H I L P K S L H G I S L H S
C Flank
4070 AAT CGG TGG AAC TGC GAC TGC CGC CTT CTA GAC ATC CAC TTC TGG CTG GTC AAC TAT
TTA GCC ACC TTG ACG CTG ACG GCG GAA GAT CTG TAG GTG AAG ACC GAC CAG TTG ATA
207 N R W N C D C R L L D I H F W L V N Y
4127 AAC ACG CCT CTG GCG GAG GAA CCC AAA TGT ATG GAA CCG GCG AGG CTG AAA GGT CAG
TTG TGC GGA GAC CGC CTC CTT GGG TTT ACA TAC CTT GGC CGC TCC GAC TTT CCA GTC
226 N T P L A E E P K C M E P A R L K G Q
4184 GTG ATC AAG AGC CTG CAG CGG GAG CAG CTG GCC TGT CTG CCG GAG GTT AGT CCC CAG
CAC TAG TTC TCG GAC GTC GCC CTC GTC GAC CGG ACA GAC GGC CTC CAA TCA GGG GTC
245 V I K S L Q R E Q L A C L P E V S P Q
4241 TCG AGT TAT ACG GAG GTG AGT GAG GGC AGG AAC ATG TCC ATC ACC TGC CTG GTC AGG
AGC TCA ATA TGC CTC CAC TCA CTC CCG TCC TTG TAC AGG TAG TGG ACG GAC CAG TCC
264 S S Y T E V S E G R N M S I T C L V R
4298 GCC ATC CCG GAG CCG AAG GTC CTT TGG CTG TTC AAT GGC CAG GTG ATG AGC AAC GAC
CGG TAG GGC CTC GGC TTC CAG GAA ACC GAC AAG TTA CCG GTC CAC TAC TCG TTG CTG
283 A I P E P K V L W L F N G Q V M S N D
4355 AGC CTG ATG GAC AAC CTG CAC ATG TAC TAC TAT ATC GAC GAG ACG ATC GGA GTA AGC
TCG GAC TAC CTG TTG GAC GTG TAC ATG ATG ATA TAG CTG CTC TGC TAG CCT CAT TCG
302 S L M D N L H M Y Y Y I D E T I G V S

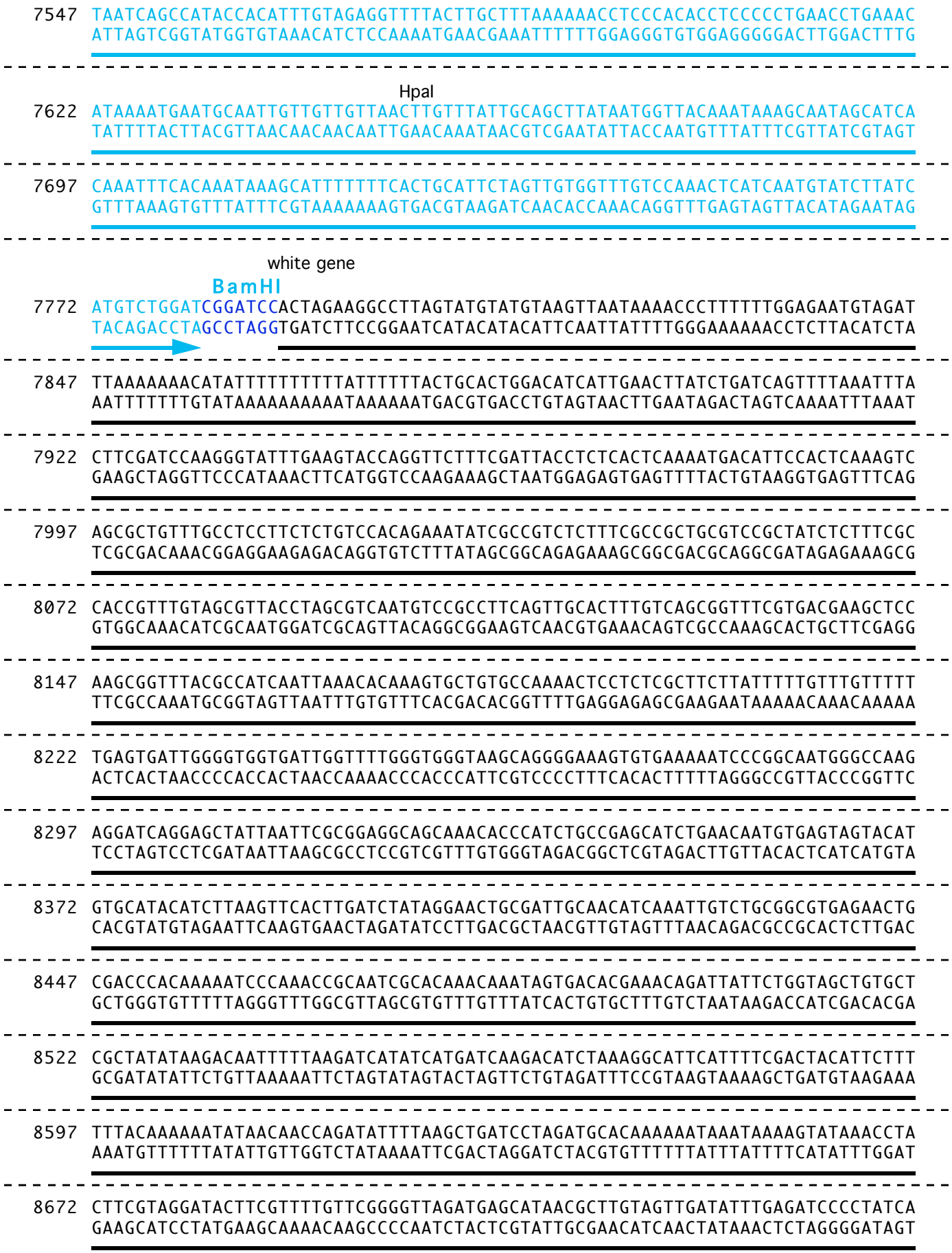

k2c t7/2 fwd oligo
4412 GGC GCC GAG GAG AAG CGC AGC GAG ATC TTC ATC TAC AAC GTT GGT GCC GAG GAT AAT
CCG CGG CTC CTC TTC GCG TCG CTC TAG AAG TAG ATG TTG CAA CCA CGG CTC CTA TTA
321 G A E E K R S E I F I Y N V G A E D N
4469 GGC ACC TTC TCC TGT GTG GGC CAG AAC ATA GCT GGC ACC ACC TTC AGT AAC TAC ACC
CCG TGG AAG AGG ACA CAC CCG GTC TTG TAT CGA CCG TGG TGG AAG TCA TTG ATG TGG
340 G T F S C V G Q N I A G T T F S N Y T
JT-CL
4526 CTG AGA GTC ATA ATC AAG GAG CCG CCG GTG GTG AAT GAG GTC TCC TTC CCC AGG GAT
GAC TCT CAG TAT TAG TTC CTC GGC GGC CAC CAC TTA CTC CAG AGG AAG GGG TCC CTA
359 L R V I I K E P P V V N E V S F P R D
4583 TAC ATG AAC TAC ATT GTG GCC AGC AGT GCC GGA GGC GGC ATT ATC TTC GTG GTA CTC
ATG TAC TTG ATG TAA CAC CGG TCG TCA CGG CCT CCG CCG TAA TAG AAG CAC CAT GAG
378 Y M N Y I V A S S A G G G I I F V V L
4640 CTC TGC ACC ATA GTG GTC AAG TGC AAG AAG ACC TCA GAG CCG GCC AAG CAG CGC AAG
GAG ACG TGG TAT CAC CAG TTC ACG TTC TTC TGG AGT CTC GGC CGG TTC GTC GCG TTC
397 L C T I V V K C K K T S E P A K Q R K
4697 AAG TGC GAT CAG GTG ACG AGT ATT GCC GGT GGC ACT GAC TCC TCG ACG GGG AGC ACC
TTC ACG CTA GTC CAC TGC TCA TAA CGG CCA CCG TGA CTG AGG AGC TGC CCC TCG TGG
416 K C D Q V T S I A G G T D S S T G S T
4754 CAG GAC ACG GGC ATG GGC ATG ATG AAG TGC GCC TCG ATA CTG AAT GAT GGC GGT GAT
GTC CTG TGC CCG TAC CCG TAC TAC TTC ACG CGG AGC TAT GAC TTA CTA CCG CCA CTA
435 Q D T G M G M M K C A S I L N D G G D
4811 AGT ATG AAC GGA AAC GCA GGA CTT CTA CTG GGC GAT ACC TTG ACA CCC ACC AAG GCG
TCA TAC TTG CCT TTG CGT CCT GAA GAT GAC CCG CTA TGG AAC TGT GGG TGG TTC CGC
454 S M N G N A G L L L G D T L T P T K A
4868 GCG AAT GGA GCA GCT GGC GGT GGC ATT ATT TTG GGC AAT CAG ATG AAG CAG AAC CTA
CGC TTA CCT CGT CGA CCG CCA CCG TAA TAA AAC CCG TTA GTC TAC TTC GTC TTG GAT
473 A N G A A G G G I I L G N Q M K Q N L
4925 CTC CTC TAC GCC ACT CCG AAC TCC GCC CAG CAG CAG CTG CAG CTG AAT GTC AAC CTG
GAG GAG ATG CGG TGA GGC TTG AGG CGG GTC GTC GTC GAC GTC GAC TTA CAG TTG GAC
492 L L Y A T P N S A Q Q Q L Q L N V N L

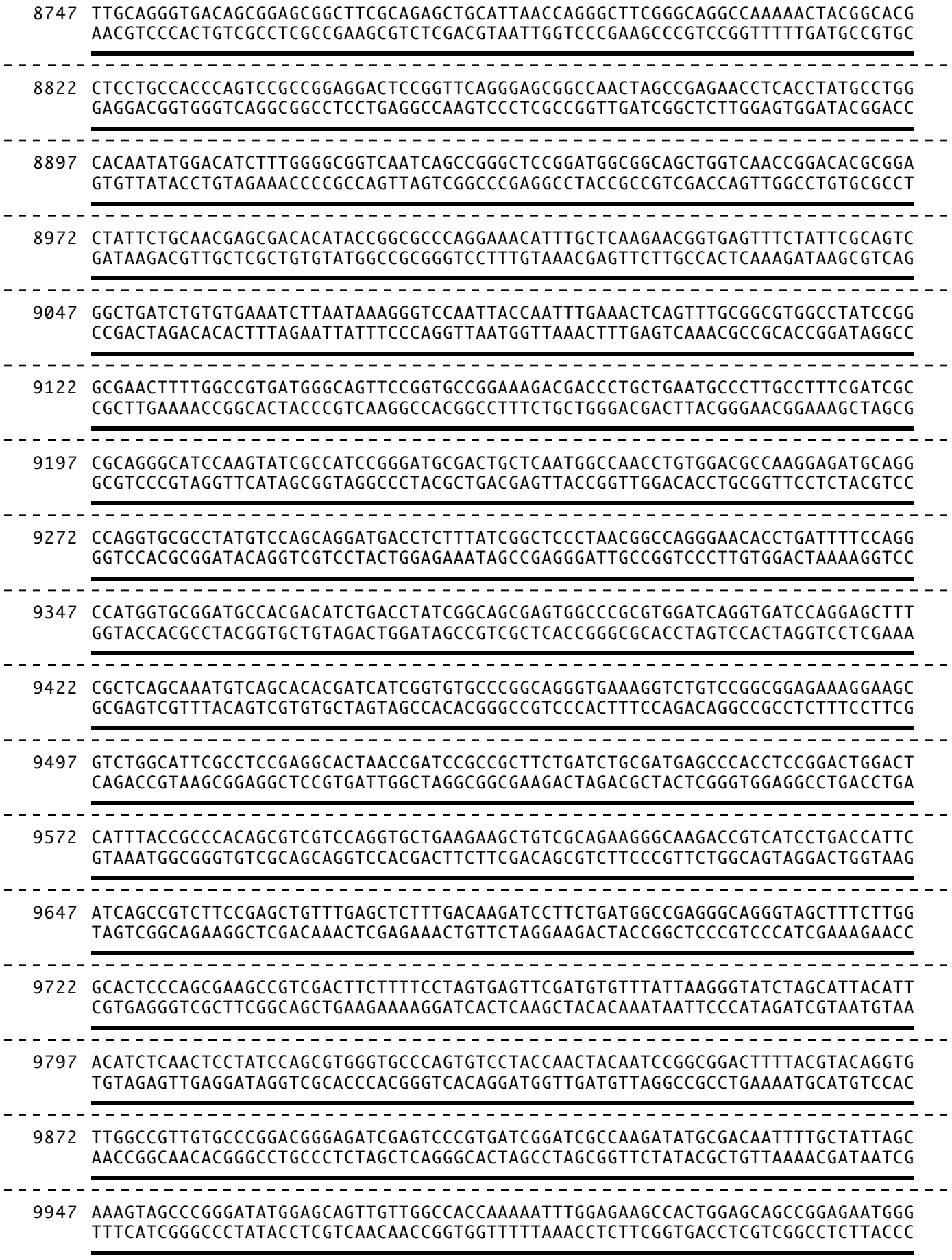

4982 ATG GGC ACT GGA CCG GGA TCA CCG CCG TTG CTC CTG AGC AAT GGC CAC GGC TTG GCG
TAC CCG TGA CCT GGC CCT AGT GGC GGC AAC GAG GAC TCG TTA CCG GTG CCG AAC CGC
511 M G T G P G S P P L L L S N G H G L A
5039 GCA GCC TAC TGC TCT CCT CCA GCT TCG CTG CGG AAC TAC CAA GAG AAA AAT CCG GAC
CGT CGG ATG ACG AGA GGA GGT CGA AGC GAC GCC TTG ATG GTT CTC TTT TTA GGC CTG
530 A A Y C S P P A S L R N Y Q E K N P D
5096 TTG GTC AAC GAT GCG GAG AGT GTC AAG CAC AAG CTT AAG ACG GCG GTA AGT CTG GAC
AAC CAG TTG CTA CGC CTC TCA CAG TTC GTG TTC GAA TTC TGC CGC CAT TCA GAC CTG
549 L V N D A E S V K H K L K T A V S L D
5153 GGA GCC GGG GAG TAC GAG ACG CAG AGC GAC TGT GGT CAG TAC GAG GGC TGC TAT CAG
CCT CGG CCC CTC ATG CTC TGC GTC TCG CTG ACA CCA GTC ATG CTC CCG ACG ATA GTC
568 G A G E Y E T Q S D C G Q Y E G C Y Q
5210 CTG GCG GCC GCT CCA CAT CCG CAT CAG GGA CAC CAG CAC CCT CAT CCG GGA CAT CCG
GAC CGC CGG CGA GGT GTA GGC GTA GTC CCT GTG GTC GTG GGA GTA GGC CCT GTA GGC
587 L A A A P H P H Q G H Q H P H P G H P
5267 CTG ATG GAA CGT TTT GCC CAG GCG ATG ACC ACT TTG CCG CGC GGC ATG CAA CTG AAG
GAC TAC CTT GCA AAA CGG GTC CGC TAC TGG TGA AAC GGC GCG CCG TAC GTT GAC TTC
606 L M E R F A Q A M T T L P R G M Q L K
5324 CCG GCT CCC CAT CAA GTT GAT GTC CAC CTG AAT CCG GTG TGC TTC CTG GGC CAA GAT
GGC CGA GGG GTA GTT CAA CTA CAG GTG GAC TTA GGC CAC ACG AAG GAC CCG GTT CTA
625 P A P H Q V D V H L N P V C F L G Q D
5381 GGA TCC TTC GCA TAT GAT TAC AGC AGT GCC CAT ATG GTG CAG CAG CCA CCT CAG CAA
CCT AGG AAG CGT ATA CTA ATG TCG TCA CGG GTA TAC CAC GTC GTC GGT GGA GTC GTT
644 G S F A Y D Y S S A H M V Q Q P P Q Q
5438 CAG CAG CAG CAG CAA CAG GTG CAG CCT GCC AAT AAC TTC TAT CGC ACG TTG CCA CAC
GTC GTC GTC GTC GTT GTC CAC GTC GGA CGG TTA TTG AAG ATA GCG TGC AAC GGT GTG
663 Q Q Q Q Q Q V Q P A N N F Y R T L P H
5495 AAT AGG TTG CAC AAA CAG CAG CAA TTT CAG GCG GCT GCG GCG GCA GGC GGA AAT GTC
TTA TCC AAC GTG TTT GTC GTC GTT AAA GTC CGC CGA CGC CGC CGT CCG CCT TTA CAG
682 N R L H K Q Q Q F Q A A A A A G G N V
5552 GGT GTG GGT GGC AAT CCC ACA CTG CGC TAC AGC CTC GAG GCC GAG TTC ATC CAG AGG
CCA CAC CCA CCG TTA GGG TGT GAC GCG ATG TCG GAG CTC CGG CTC AAG TAG GTC TCC
701 G V G G N P T L R Y S L E A E F I Q R

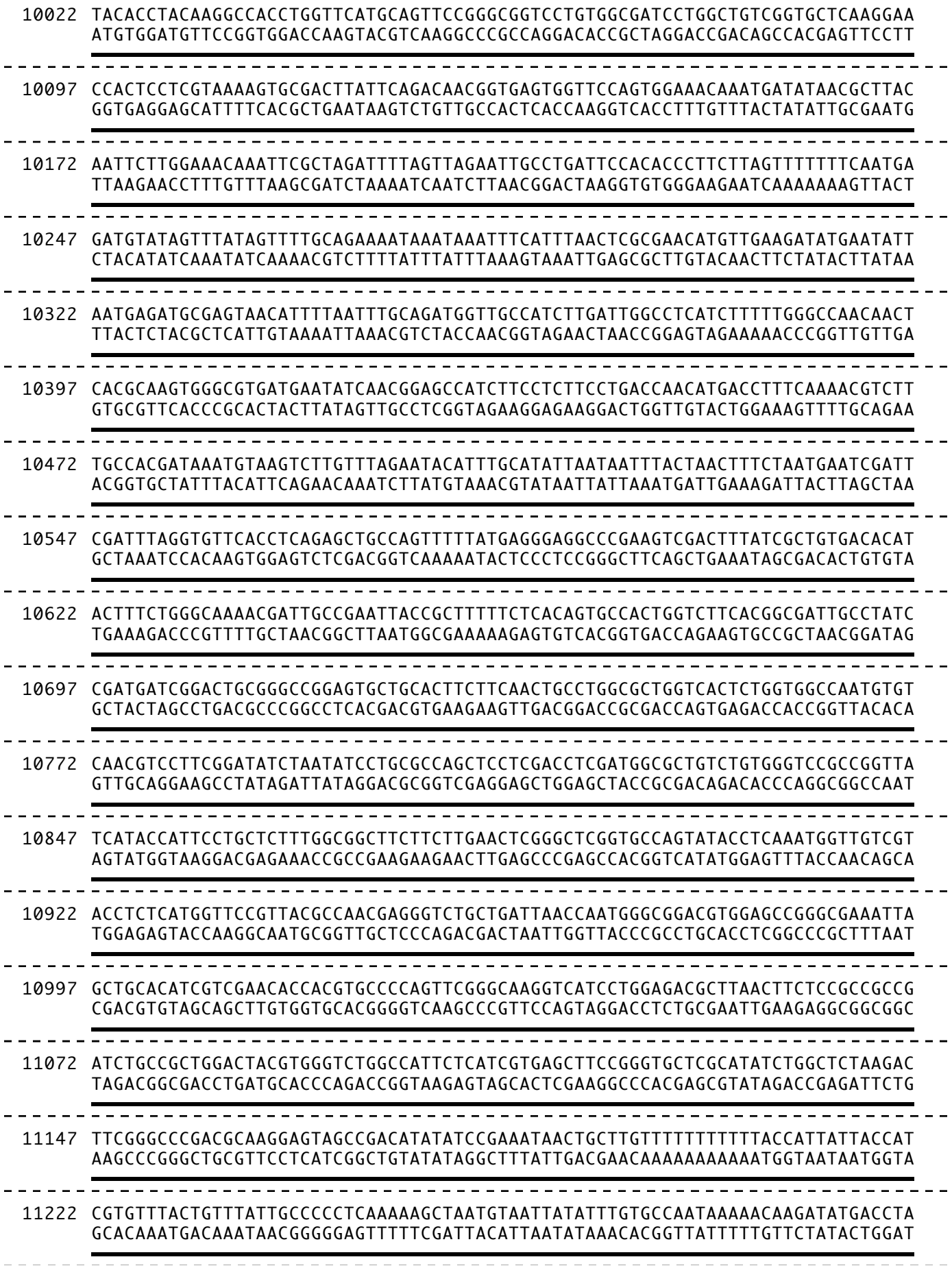

5609 GGT CCG ACG GTG AGC TAC GAG AAG TAC CAG CTG CCC AAT GTG CGC TTC ACA GCG GAG
CCA GGC TGC CAC TCG ATG CTC TTC ATG GTC GAC GGG TTA CAC GCG AAG TGT CGC CTC
720 G P T V S Y E K Y Q L P N V R F T A E
5666 GGT TAT CCT CAG CAG CAG CAG CAG CAA CAG CAG CAA CTT CAG CAG CAG CAA CAG TTG
CCA ATA GGA GTC GTC GTC GTC GTC GTT GTC GTC GTT GAA GTC GTC GTC GTT GTC AAC
739 G Y P Q Q Q Q Q Q Q Q Q L Q Q Q Q Q L
5723 CAG CTC CAG CAG CAG CAT CAG TTC CCC TCA CCG CCA GAG GGC TAC AAA AGC
GTC GAG GTC GTC GTC GTA GTC AAG GGG AGT GGC GGT CTC CCG ATG TTT TCG
758 Q L Q Q Q H Q F P S P P E G Y K S
5774 GAT TTG GCG GTG ATG CCG GCG CCG TTT CAG CAA TGG CCC AGC TGT TTG
CTA AAC CGC CAC TAC GGC CGC GGC AAA GTC GTT ACC GGG TCG ACA AAC
775 D L A V M P A P F Q Q W P S C L
5822 CCC GGC TAC CGA TTC GCC CAG TCA CCC ACC AGC CTG CCA GCG GTT GCA
GGG CCG ATG GCT AAG CGG GTC AGT GGG TGG TCG GAC GGT CGC CAA CGT
791 P G Y R F A Q S P T S L P A V A
RNAi 42449
Oligo W66
5870 ACT CCA CCA CCA GCC GCA GTG GTG GCA ACA CCA CCA CCG CCC ACA TCA
TGA GGT GGT GGT CGG CGT CAC CAC CGT TGT GGT GGT GGC GGG TGT AGT
807 T P P P A A V V A T P P P P T S
5918 GCA GTA AGC ACA CAA TCC ACG GCC ACA TCC ACC ATT CCG GAG CTG GAT
CGT CAT TCG TGT GTT AGG TGC CGG TGT AGG TGG TAA GGC CTC GAC CTA
823 A V S T Q S T A T S T I P E L D
5966 GAA AGC GAG GCG AGT TCG CCG CGC CTC GAG GAA GCC GCC GGG TCT GCA
CTT TCG CTC CGC TCA AGC GGC GCG GAG CTC CTT CGG CGG CCC AGA CGT
839 E S E A S S P R L E E A A G S A
6014 GCG CCA CCT GCC GGT GAG GAG GAG AGT TCG GAC ACC GCA AAA CTC AAA
CGC GGT GGA CGG CCA CTC CTC CTC TCA AGC CTG TGG CGT TTT GAG TTT
855 A P P A G E E E S S D T A K L K
6062 CAG CTT AAT GGC CCC TTG GCC GAC AGT CCC GAC GAG GGA TAC GTG GGC
GTC GAA TTA CCG GGG AAC CGG CTG TCA GGG CTG CTC CCT ATG CAC CCG
871 Q L N G P L A D S P D E G Y V G

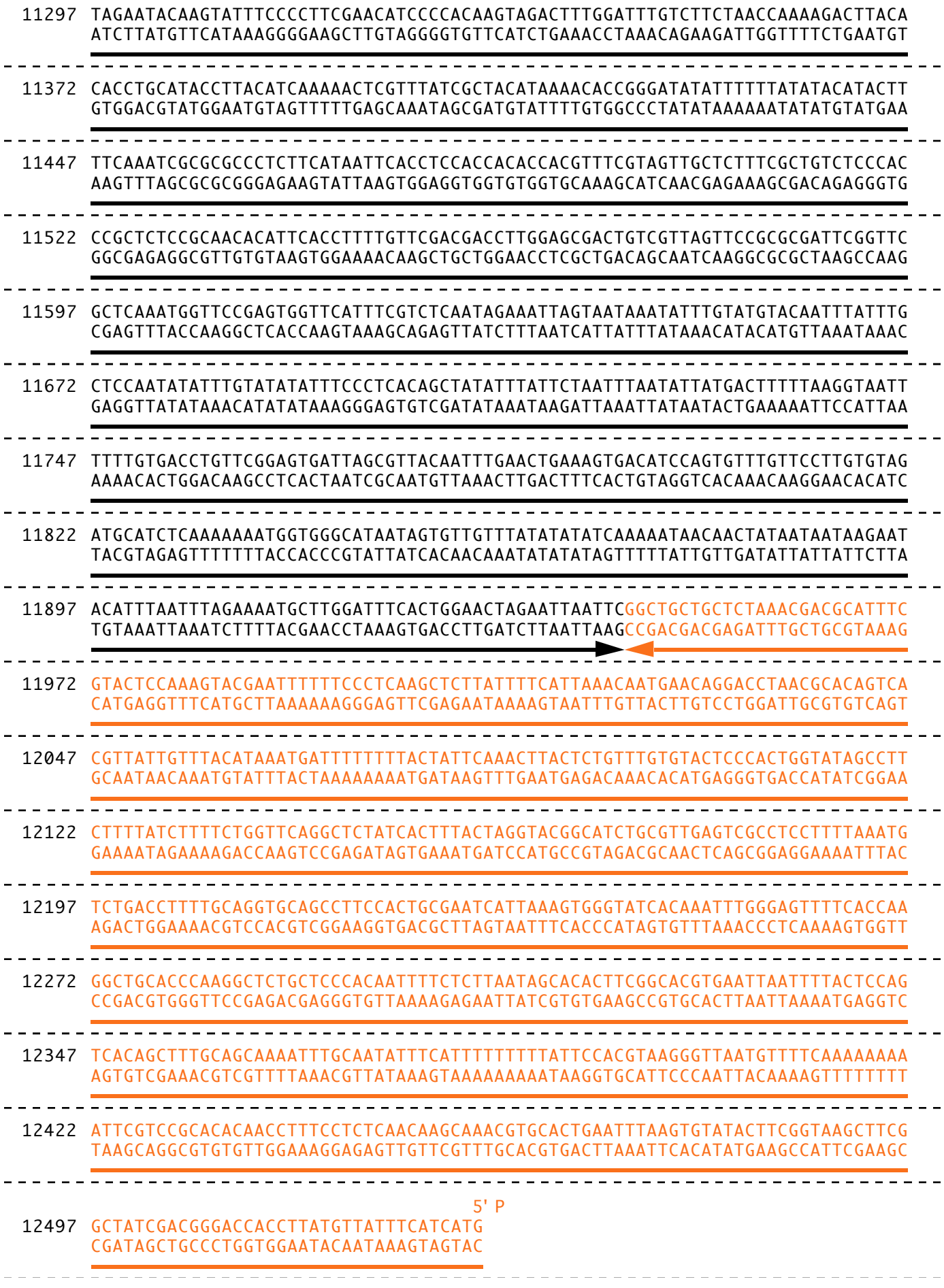

attB2
6110 GAT GGC CAG GAA ACC AGC GAC ATT GAC CCA GCT TTC TTG TAC AAA GTG
CTA CCG GTC CTT TGG TCG CTG TAA CTG GGT CGA AAG AAC ATG TTT CAC
887 D G Q E T S D I D P A F L Y K V
KpnI
6158 GTG GTA CCG CGG GCC CGG GAT CCA CCG GTC GCC ACC ATG GTG AGC AAG
CAC CAT GGC GCC CGG GCC CTA GGT GGC CAG CGG TGG TAC CAC TCG TTC
903 V V P R A R D P P V A T M V S K
6206 GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC
CCG CTC CTC GAC AAG TGG CCC CAC CAC GGG TAG GAC CAG CTC GAC CTG CCG CTG
919 G E E L F T G V V P I L V E L D G D
6260 GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC GGC
CAT TTG CCG GTG TTC AAG TCG CAC AGG CCG CTC CCG CTC CCG CTA CGG TGG ATG CCG
937 V N G H K F S V S G E G E G D A T Y G
6317 AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC
TTC GAC TGG GAC TTC AAG TAG ACG TGG TGG CCG TTC GAC GGG CAC GGG ACC GGG TGG
956 K L T L K F I C T T G K L P V P W P T
6374 CTC GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG
GAG CAC TGG TGG GAC TGG ATG CCG CAC GTC ACG AAG TCG GCG ATG GGG CTG GTG TAC
975 L V T T L T Y G V Q C F S R Y P D H M
6431 AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG CGC ACC
TTC GTC GTG CTG AAG AAG TTC AGG CGG TAC GGG CTT CCG ATG CAG GTC CTC GCG TGG
994 K Q H D F F K S A M P E G Y V Q E R T
6488 ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC
TAG AAG AAG TTC CTG CTG CCG TTG ATG TTC TGG GCG CGG CTC CAC TTC AAG CTC CCG
1013 I F F K D D G N Y K T R A E V K F E G
6545 GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC
CTG TGG GAC CAC TTG GCG TAG CTC GAC TTC CCG TAG CTG AAG TTC CTC CTG CCG TTG
1032 D T L V N R I E L K G I D F K E D G N
6602 ATC CTG GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC
TAG GAC CCC GTG TTC GAC CTC ATG TTG ATG TTG TCG GTG TTG CAG ATA TAG TAC CGG
1051 I L G H K L E Y N Y N S H N V Y I M A

```
  6659 GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC
       CTG TTC GTC TTC TTG CCG TAG TTC CAC TTG AAG TTC TAG GCG GTG TTG TAG CTC CTG

  1070  D   K   Q   K   N   G   I   K   V   N   F   K   I   R   H   N   I   E   D

  6716 GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC CCC
       CCG TCG CAC GTC GAG CGG CTG GTG ATG GTC GTC TTG TGG GGG TAG CCG CTG CCG GGG

  1089  G   S   V   Q   L   A   D   H   Y   Q   Q   N   T   P   I   G   D   G   P

  6773 GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC
       CAC GAC GAC GGG CTG TTG GTG ATG GAC TCG TGG GTC AGG CGG GAC TCG TTT CTG GGG

  1108  V   L   L   P   D   N   H   Y   L   S   T   Q   S   A   L   S   K   D   P

  6830 AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT
       TTG CTC TTC GCG CTA GTG TAC CAG GAC GAC CTC AAG CAC TGG CGG CGG CCC TAG TGA

  1127  N   E   K   R   D   H   M   V   L   L   E   F   V   T   A   A   G   I   T

                                                                  SV40 Poly A
                                                 NotI        XbaI  BglII
  6887 CTC GGC ATG GAC GAG CTG TAC AAG TAA AGC GGC CGC GAC TCT AGA GATCTTTGTGAAGGA
       GAG CCG TAC CTG CTC GAC ATG TTC ATT TCG CCG GCG CTG AGA TCT CTAGAAACACTTCCT

  1146  L   G   M   D   E   L   Y   K   •   S   G   R

  6947 ACCTTACTTCTGTGGTGTGACATAATTGGACAAACTACCTACAGAGATTTAAAGCTCTAAGGTAAATATAAAATT
       TGGAATGAAGACACCACACTGTATTAACCTGTTTGATGGATGTCTCTAAATTTCGAGATTCCATTTATATTTTAA

  7022 TTTAAGTGTATAATGTGTTAAACTACTGATTCTAATTGTTTGTGTATTTTAGATTCCAACCTATGGAACTGATGA
       AAATTCACATATTACACAATTTGATGACTAAGATTAACAAACACATAAAATCTAAGGTTGGATACCTTGACTACT

  7097 ATGGGAGCAGTGGTGGAATGCCTTTAATGAGGAAAACCTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGATGA
       TACCCTCGTCACCACCTTACGGAAATTACTCCTTTTGGACAAAACGAGTCTTCTTTACGGTAGATCACTACTACT

  7172 GGCTACTGCTGACTCTCAACATTCTACTCCTCCAAAAAAGAAGAGAAAGGTAGAAGACCCCAAGGACTTTCCTTC
       CCGATGACGACTGAGAGTTGTAAGATGAGGAGGTTTTTTCTTCTCTTTCCATCTTCTGGGGTTCCTGAAAGGAAG

  7247 AGAATTGCTAAGTTTTTTGAGTCATGCTGTGTTTAGTAATAGAACTCTTGCTTGCTTTGCTATTTACACCACAAA
       TCTTAACGATTCAAAAAACTCAGTACGACACAAATCATTATCTTGAGAACGAACGAAACGATAAATGTGGTGTTT

  7322 GGAAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATATTCTGTAACCTTTATAAGTAGGCATAACAGTTA
       CCTTTTTCGACGTGACGATATGTTCTTTTAATACCTTTTTATAAGACATTGGAAATATTCATCCGTATTGTCAAT

  7397 TAATCATAACATACTGTTTTTTCTTACTCCACACAGGCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAATT
       ATTAGTATTGTATGACAAAAAAGAATGAGGTGTGTCCGTATCTCACAGACGATAATTATTGATACGAGTTTTTAA

  7472 GTGTACCTTTAGCTTTTTAATTTGTAAAGGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAGATCA
       CACATGGAAATCGAAAAATTAAACATTTCCCCAATTATTCCTTATAAACTACATATCACGGAACTGATCTCTAGT
```


7547
TAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAAC
ATTAGTCGGTATGGTGTAAACATCTCCAAAATGAACGAAATTTTTTGGAGGGTGTGGAGGGGGACTTGGACTTTG
HpaI
7622
ATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA
TATTTTACTTACGTTAACAACAACAATTGAACAAATAACGTCGAATATTACCAATGTTTATTTCGTTATCGTAGT
7697
CAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATC
GTTTAAAGTGTTTATTTCGTAAAAAAAGTGACGTAAGATCAACACCAAACAGGTTTGAGTAGTTACATAGAATAG
white gene
BamHI
7772
ATGTCTGGATCGGATCCACTAGAAGGCCTTAGTATGTATGTAAGTTAATAAAACCCTTTTTTGGAGAATGTAGAT
TACAGACCTAGCCTAGGTGATCTTCCGGAATCATACATACATTCAATTATTTTGGGAAAAAACCTCTTACATCTA
7847
TTAAAAAAACATATTTTTTTTTTATTTTTTACTGCACTGGACATCATTGAACTTATCTGATCAGTTTTAAATTTA
AATTTTTTTGTATAAAAAAAAAATAAAAAATGACGTGACCTGTAGTAACTTGAATAGACTAGTCAAAATTTAAAT
7922
CTTCGATCCAAGGGTATTTGAAGTACCAGGTTCTTTCGATTACCTCTCACTCAAAATGACATTCCACTCAAAGTC
GAAGCTAGGTTCCCATAAACTTCATGGTCCAAGAAAGCTAATGGAGAGTGAGTTTTACTGTAAGGTGAGTTTCAG
7997
AGCGCTGTTTGCCTCCTTCTCTGTCCACAGAAATATCGCCGTCTCTTTCGCCGCTGCGTCCGCTATCTCTTTCGC
TCGCGACAAACGGAGGAAGAGACAGGTGTCTTTATAGCGGCAGAGAAAGCGGCGACGCAGGCGATAGAGAAAGCG
8072
CACCGTTTGTAGCGTTACCTAGCGTCAATGTCCGCCTTCAGTTGCACTTTGTCAGCGGTTTCGTGACGAAGCTCC
GTGGCAAACATCGCAATGGATCGCAGTTACAGGCGGAAGTCAACGTGAAACAGTCGCCAAAGCACTGCTTCGAGG
8147
AAGCGGTTTACGCCATCAATTAAACACAAAGTGCTGTGCCAAAACTCCTCTCGCTTCTTATTTTTGTTTGTTTTT
TTCGCCAAATGCGGTAGTTAATTTGTGTTTCACGACACGGTTTTGAGGAGAGCGAAGAATAAAAACAAACAAAAA
8222
TGAGTGATTGGGGTGGTGATTGGTTTTGGGTGGGTAAGCAGGGGAAAGTGTGAAAAATCCCGGCAATGGGCCAAG
ACTCACTAACCCCACCACTAACCAAAACCCACCCATTCGTCCCCTTTCACACTTTTTAGGGCCGTTACCCGGTTC
8297
AGGATCAGGAGCTATTAATTCGCGGAGGCAGCAAACACCCATCTGCCGAGCATCTGAACAATGTGAGTAGTACAT
TCCTAGTCCTCGATAATTAAGCGCCTCCGTCGTTTGTGGGTAGACGGCTCGTAGACTTGTTACACTCATCATGTA
8372
GTGCATACATCTTAAGTTCACTTGATCTATAGGAACTGCGATTGCAACATCAAATTGTCTGCGGCGTGAGAACTG
CACGTATGTAGAATTCAAGTGAACTAGATATCCTTGACGCTAACGTTGTAGTTTAACAGACGCCGCACTCTTGAC
8447
CGACCCACAAAAATCCCAAACCGCAATCGCACAAACAAATAGTGACACGAAACAGATTATTCTGGTAGCTGTGCT
GCTGGGTGTTTTTAGGGTTTGGCGTTAGCGTGTTTGTTTATCACTGTGCTTTGTCTAATAAGACCATCGACACGA
8522
CGCTATATAAGACAATTTTTAAGATCATATCATGATCAAGACATCTAAAGGCATTCATTTTCGACTACATTCTTT
GCGATATATTCTGTTAAAAATTCTAGTATAGTACTAGTTCTGTAGATTTCCGTAAGTAAAAGCTGATGTAAGAAA
8597
TTTACAAAAAATATAACAACCAGATATTTTAAGCTGATCCTAGATGCACAAAAAATAAATAAAAGTATAAACCTA
AAATGTTTTTTATATTGTTGGTCTATAAAATTCGACTAGGATCTACGTGTTTTTTATTTATTTTCATATTTGGAT
8672
CTTCGTAGGATACTTCGTTTTGTTCGGGGTTAGATGAGCATAACGCTTGTAGTTGATATTTGAGATCCCCTATCA
GAAGCATCCTATGAAGCAAAACAAGCCCCAATCTACTCGTATTGCGAACATCAACTATAAACTCTAGGGGATAGT


8747 TTGCAGGGTGACAGCGGAGCGGCTTCGCAGAGCTGCATTAACCAGGGCTTCGGGCAGGCCAAAAACTACGGCACG
AACGTCCCACTGTCGCCTCGCCGAAGCGTCTCGACGTAATTGGTCCCGAAGCCCGTCCGGTTTTTGATGCCGTGC
8822 CTCCTGCCACCCAGTCCGCCGGAGGACTCCGGTTCAGGGAGCGGCCAACTAGCCGAGAACCTCACCTATGCCTGG
GAGGACGGTGGGTCAGGCGGCCTCCTGAGGCCAAGTCCCTCGCCGGTTGATCGGCTCTTGGAGTGGATACGGACC
8897 CACAATATGGACATCTTTGGGGCGGTCAATCAGCCGGGCTCCGGATGGCGGCAGCTGGTCAACCGGACACGCGGA
GTGTTATACCTGTAGAAACCCCGCCAGTTAGTCGGCCCGAGGCCTACCGCCGTCGACCAGTTGGCCTGTGCGCCT
8972 CTATTCTGCAACGAGCGACACATACCGGCGCCCAGGAAACATTTGCTCAAGAACGGTGAGTTTCTATTCGCAGTC
GATAAGACGTTGCTCGCTGTGTATGGCCGCGGGTCCTTTGTAAACGAGTTCTTGCCACTCAAAGATAAGCGTCAG
9047 GGCTGATCTGTGTGAAATCTTAATAAAGGGTCCAATTACCAATTTGAAACTCAGTTTGCGGCGTGGCCTATCCGG
CCGACTAGACACACTTTAGAATTATTTCCCAGGTTAATGGTTAAACTTTGAGTCAAACGCCGCACCGGATAGGCC
9122 GCGAACTTTTGGCCGTGATGGGCAGTTCCGGTGCCGGAAAGACGACCCTGCTGAATGCCCTTGCCTTTCGATCGC
CGCTTGAAAACCGGCACTACCCGTCAAGGCCACGGCCTTTCTGCTGGGACGACTTACGGGAACGGAAAGCTAGCG
9197 CGCAGGGCATCCAAGTATCGCCATCCGGGATGCGACTGCTCAATGGCCAACCTGTGGACGCCAAGGAGATGCAGG
GCGTCCCGTAGGTTCATAGCGGTAGGCCCTACGCTGACGAGTTACCGGTTGGACACCTGCGGTTCCTCTACGTCC
9272 CCAGGTGCGCCTATGTCCAGCAGGATGACCTCTTTATCGGCTCCCTAACGGCCAGGGAACACCTGATTTTCCAGG
GGTCCACGCGGATACAGGTCGTCCTACTGGAGAAATAGCCGAGGGATTGCCGGTCCCTTGTGGACTAAAAGGTCC
9347 CCATGGTGCGGATGCCACGACATCTGACCTATCGGCAGCGAGTGGCCCGCGTGGATCAGGTGATCCAGGAGCTTT
GGTACCACGCCTACGGTGCTGTAGACTGGATAGCCGTCGCTCACCGGGCGCACCTAGTCCACTAGGTCCTCGAAA
9422 CGCTCAGCAAATGTCAGCACACGATCATCGGTGTGCCCGGCAGGGTGAAAGGTCTGTCCGGCGGAGAAAGGAAGC
GCGAGTCGTTTACAGTCGTGTGCTAGTAGCCACACGGGCCGTCCCACTTTCCAGACAGGCCGCCTCTTTCCTTCG
9497 GTCTGGCATTCGCCTCCGAGGCACTAACCGATCCGCCGCTTCTGATCTGCGATGAGCCCACCTCCGGACTGGACT
CAGACCGTAAGCGGAGGCTCCGTGATTGGCTAGGCGGCGAAGACTAGACGCTACTCGGGTGGAGGCCTGACCTGA
9572 CATTTACCGCCCACAGCGTCGTCCAGGTGCTGAAGAAGCTGTCGCAGAAGGGCAAGACCGTCATCCTGACCATTC
GTAAATGGCGGGTGTCGCAGCAGGTCCACGACTTCTTCGACAGCGTCTTCCCGTTCTGGCAGTAGGACTGGTAAG
9647 ATCAGCCGTCTTCCGAGCTGTTTGAGCTCTTTGACAAGATCCTTCTGATGGCCGAGGGCAGGGTAGCTTTCTTGG
TAGTCGGCAGAAGGCTCGACAAACTCGAGAAACTGTTCTAGGAAGACTACCGGCTCCCGTCCCATCGAAAGAACC
9722 GCACTCCCAGCGAAGCCGTCGACTTCTTTTCCTAGTGAGTTCGATGTGTTTATTAAGGGTATCTAGCATTACATT
CGTGAGGGTCGCTTCGGCAGCTGAAGAAAAGGATCACTCAAGCTACACAAATAATTCCCATAGATCGTAATGTAA
9797 ACATCTCAACTCCTATCCAGCGTGGGTGCCCAGTGTCCTACCAACTACAATCCGGCGGACTTTTACGTACAGGTG
TGTAGAGTTGAGGATAGGTCGCACCCACGGGTCACAGGATGGTTGATGTTAGGCCGCCTGAAAATGCATGTCCAC
9872 TTGGCCGTTGTGCCCGGACGGGAGATCGAGTCCCGTGATCGGATCGCCAAGATATGCGACAATTTTGCTATTAGC
AACCGGCAACACGGGCCTGCCCTCTAGCTCAGGGCACTAGCCTAGCGGTTCTATACGCTGTTAAAACGATAATCG
9947 AAAGTAGCCCGGGATATGGAGCAGTTGTTGGCCACCAAAAATTTGGAGAAGCCACTGGAGCAGCCGGAGAATGGG
TTTCATCGGGCCCTATACCTCGTCAACAACCGGTGGTTTTTAAACCTCTTCGGTGACCTCGTCGGCCTCTTACCC


10022 TACACCTACAAGGCCACCTGGTTCATGCAGTTCCGGGCGGTCCTGTGGCGATCCTGGCTGTCGGTGCTCAAGGAA
ATGTGGATGTTCCGGTGGACCAAGTACGTCAAGGCCCGCCAGGACACCGCTAGGACCGACAGCCACGAGTTCCTT
10097 CCACTCCTCGTAAAAGTGCGACTTATTCAGACAACGGTGAGTGGTTCCAGTGGAAACAAATGATATAACGCTTAC
GGTGAGGAGCATTTTCACGCTGAATAAGTCTGTTGCCACTCACCAAGGTCACCTTTGTTTACTATATTGCGAATG
10172 AATTCTTGGAAACAAATTCGCTAGATTTTAGTTAGAATTGCCTGATTCCACACCCTTCTTAGTTTTTTTCAATGA
TTAAGAACCTTTGTTTAAGCGATCTAAAATCAATCTTAACGGACTAAGGTGTGGGAAGAATCAAAAAAAGTTACT
10247 GATGTATAGTTTATAGTTTTGCAGAAAATAAATAAATTTCATTTAACTCGCGAACATGTTGAAGATATGAATATT
CTACATATCAAATATCAAAACGTCTTTTATTTATTTAAAGTAAATTGAGCGCTTGTACAACTTCTATACTTATAA
10322 AATGAGATGCGAGTAACATTTTAATTTGCAGATGGTTGCCATCTTGATTGGCCTCATCTTTTTGGGCCAACAACT
TTACTCTACGCTCATTGTAAAATTAAACGTCTACCAACGGTAGAACTAACCGGAGTAGAAAAACCCGGTTGTTGA
10397 CACGCAAGTGGGCGTGATGAATATCAACGGAGCCATCTTCCTCTTCCTGACCAACATGACCTTTCAAAACGTCTT
GTGCGTTCACCCGCACTACTTATAGTTGCCTCGGTAGAAGGAGAAGGACTGGTTGTACTGGAAAGTTTTGCAGAA
10472 TGCCACGATAAATGTAAGTCTTGTTTAGAATACATTTGCATATTAATAATTTACTAACTTTCTAATGAATCGATT
ACGGTGCTATTTACATTCAGAACAAATCTTATGTAAACGTATAATTATTAAATGATTGAAAGATTACTTAGCTAA
10547 CGATTTAGGTGTTCACCTCAGAGCTGCCAGTTTTTATGAGGGAGGCCCGAAGTCGACTTTATCGCTGTGACACAT
GCTAAATCCACAAGTGGAGTCTCGACGGTCAAAAATACTCCCTCCGGGCTTCAGCTGAAATAGCGACACTGTGTA
10622 ACTTTCTGGGCAAAACGATTGCCGAATTACCGCTTTTTCTCACAGTGCCACTGGTCTTCACGGCGATTGCCTATC
TGAAAGACCCGTTTTGCTAACGGCTTAATGGCGAAAAAGAGTGTCACGGTGACCAGAAGTGCCGCTAACGGATAG
10697 CGATGATCGGACTGCGGGCCGGAGTGCTGCACTTCTTCAACTGCCTGGCGCTGGTCACTCTGGTGGCCAATGTGT
GCTACTAGCCTGACGCCCGGCCTCACGACGTGAAGAAGTTGACGGACCGCGACCAGTGAGACCACCGGTTACACA
10772 CAACGTCCTTCGGATATCTAATATCCTGCGCCAGCTCCTCGACCTCGATGGCGCTGTCTGTGGGTCCGCCGGTTA
GTTGCAGGAAGCCTATAGATTATAGGACGCGGTCGAGGAGCTGGAGCTACCGCGACAGACACCCAGGCGGCCAAT
10847 TCATACCATTCCTGCTCTTTGGCGGCTTCTTCTTGAACTCGGGCTCGGTGCCAGTATACCTCAAATGGTTGTCGT
AGTATGGTAAGGACGAGAAACCGCCGAAGAAGAACTTGAGCCCGAGCCACGGTCATATGGAGTTTACCAACAGCA
10922 ACCTCTCATGGTTCCGTTACGCCAACGAGGGTCTGCTGATTAACCAATGGGCGGACGTGGAGCCGGGCGAAATTA
TGGAGAGTACCAAGGCAATGCGGTTGCTCCCAGACGACTAATTGGTTACCCGCCTGCACCTCGGCCCGCTTTAAT
10997 GCTGCACATCGTCGAACACCACGTGCCCCAGTTCGGGCAAGGTCATCCTGGAGACGCTTAACTTCTCCGCCGCCG
CGACGTGTAGCAGCTTGTGGTGCACGGGGTCAAGCCCGTTCCAGTAGGACCTCTGCGAATTGAAGAGGCGGCGGC
11072 ATCTGCCGCTGGACTACGTGGGTCTGGCCATTCTCATCGTGAGCTTCCGGGTGCTCGCATATCTGGCTCTAAGAC
TAGACGGCGACCTGATGCACCCAGACCGGTAAGAGTAGCACTCGAAGGCCCACGAGCGTATAGACCGAGATTCTG
11147 TTCGGGCCCGACGCAAGGAGTAGCCGACATATATCCGAAATAACTGCTTGTTTTTTTTTTTACCATTATTACCAT
AAGCCCGGGCTGCGTTCCTCATCGGCTGTATATAGGCTTTATTGACGAACAAAAAAAAAAATGGTAATAATGGTA
11222 CGTGTTTACTGTTTATTGCCCCCTCAAAAAGCTAATGTAATTATATTTGTGCCAATAAAAACAAGATATGACCTA
GCACAAATGACAAATAACGGGGGAGTTTTTCGATTACATTAATATAAACACGGTTATTTTTGTTCTATACTGGAT

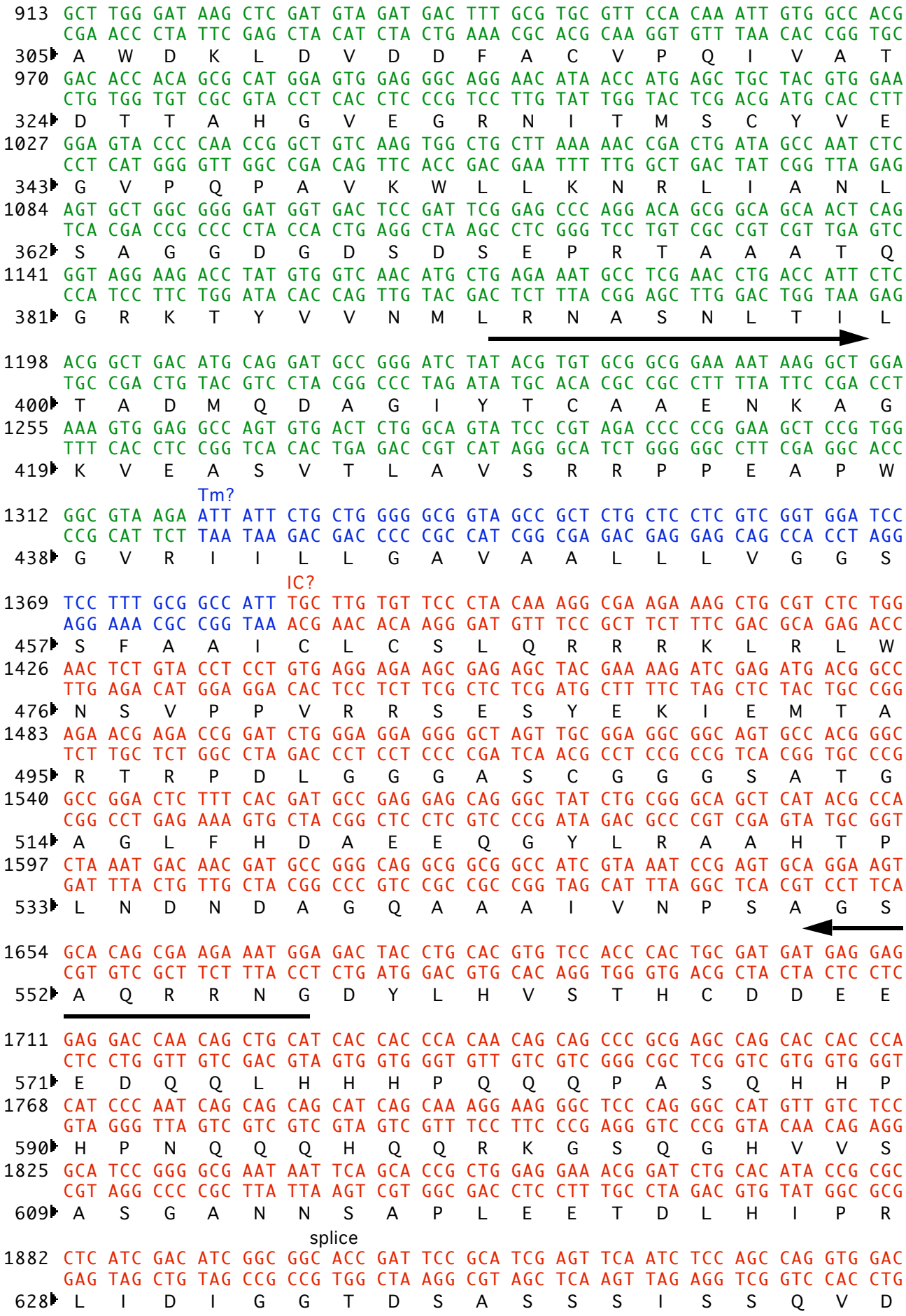

11297 TAGAATACAAGTATTTCCCCTTCGAACATCCCCACAAGTAGACTTTGGATTTGTCTTCTAACCAAAAGACTTACA
ATCTTATGTTCATAAAGGGGAAGCTTGTAGGGGTGTTCATCTGAAACCTAAACAGAAGATTGGTTTTCTGAATGT
11372 CACCTGCATACCTTACATCAAAAACTCGTTTATCGCTACATAAAACACCGGGATATATTTTTTATATACATACTT
GTGGACGTATGGAATGTAGTTTTTGAGCAAATAGCGATGTATTTTGTGGCCCTATATAAAAAATATATGTATGAA
11447 TTCAAATCGCGCGCCCTCTTCATAATTCACCTCCACCACACCACGTTTCGTAGTTGCTCTTTCGCTGTCTCCCAC
AAGTTTAGCGCGCGGGAGAAGTATTAAGTGGAGGTGGTGTGGTGCAAAGCATCAACGAGAAAGCGACAGAGGGTG
11522 CCGCTCTCCGCAACACATTCACCTTTTGTTCGACGACCTTGGAGCGACTGTCGTTAGTTCCGCGCGATTCGGTTC
GGCGAGAGGCGTTGTGTAAGTGGAAAACAAGCTGCTGGAACCTCGCTGACAGCAATCAAGGCGCGCTAAGCCAAG
11597 GCTCAAATGGTTCCGAGTGGTTCATTTCGTCTCAATAGAAATTAGTAATAAATATTTGTATGTACAATTTATTTG
CGAGTTTACCAAGGCTCACCAAGTAAAGCAGAGTTATCTTTAATCATTATTTATAAACATACATGTTAAATAAAC
11672 CTCCAATATATTTGTATATATTTCCCTCACAGCTATATTTATTCTAATTTAATATTATGACTTTTTAAGGTAATT
GAGGTTATATAAACATATATAAAGGGAGTGTCGATATAAATAAGATTAAATTATAATACTGAAAAATTCCATTAA
11747 TTTTGTGACCTGTTCGGAGTGATTAGCGTTACAATTTGAACTGAAAGTGACATCCAGTGTTTGTTCCTTGTGTAG
AAAACACTGGACAAGCCTCACTAATCGCAATGTTAAACTTGACTTTCACTGTAGGTCACAAACAAGGAACACATC
11822 ATGCATCTCAAAAAAATGGTGGGCATAATAGTGTTGTTTATATATATCAAAAATAACAACTATAATAATAAGAAT
TACGTAGAGTTTTTTTACCACCCGTATTATCACAACAAATATATATAGTTTTTATTGTTGATATTATTATTCTTA
11897 ACATTTAATTTAGAAAATGCTTGGATTTCACTGGAACTAGAATTAATTCGGCTGCTGCTCTAAACGACGCATTTC
TGTAAATTAAATCTTTTACGAACCTAAAGTGACCTTGATCTTAATTAAGCCGACGACGAGATTTGCTGCGTAAAG
11972 GTACTCCAAAGTACGAATTTTTTCCCTCAAGCTCTTATTTTCATTAAACAATGAACAGGACCTAACGCACAGTCA
CATGAGGTTTCATGCTTAAAAAAGGGAGTTCGAGAATAAAAGTAATTTGTTACTTGTCCTGGATTGCGTGTCAGT
12047 CGTTATTGTTTACATAAATGATTTTTTTTACTATTCAAACTTACTCTGTTTGTGTACTCCCACTGGTATAGCCTT
GCAATAACAAATGTATTTACTAAAAAAAAATGATAAGTTTGAATGAGACAAACACATGAGGGTGACCATATCGGAA
12122 CTTTTATCTTTTCTGGTTCAGGCTCTATCACTTTACTAGGTACGGCATCTGCGTTGAGTCGCCTCCTTTTAAATG
GAAAATAGAAAAGACCAAGTCCGAGATAGTGAAATGATCCATGCCGTAGACGCAACTCAGCGGAGGAAAATTTAC
12197 TCTGACCTTTTGCAGGTGCAGCCTTCCACTGCGAATCATTAAAGTGGGTATCACAAATTTGGGAGTTTTCACCAA
AGACTGGAAAACGTCCACGTCGGAAGGTGACGCTTAGTAATTTCACCCATAGTGTTTAAACCCTCAAAAGTGGTT
12272 GGCTGCACCCAAGGCTCTGCTCCCACAATTTTCTCTTAATAGCACACTTCGGCACGTGAATTAATTTTACTCCAG
CCGACGTGGGTTCCGAGACGAGGGTGTTAAAAGAGAATTATCGTGTGAAGCCGTGCACTTAATTAAAATGAGGTC
12347 TCACAGCTTTGCAGCAAAATTTGCAATATTTCATTTTTTTTTATTCCACGTAAGGGTTAATGTTTTCAAAAAAAA
AGTGTCGAAACGTCGTTTTAAACGTTATAAAGTAAAAAAAAATAAGGTGCATTCCCAATTACAAAAGTTTTTTTT
12422 ATTCGTCCGCACACAACCTTTCCTCTCAACAAGCAAACGTGCACTGAATTTAAGTGTATACTTCGGTAAGCTTCG
TAAGCAGGCGTGTGTTGGAAAGGAGAGTTGTTCGTTTGCACGTGACTTAAATTCACATATGAAGCCATTCGAAGC
5' P
12497 GCTATCGACGGGACCACCTTATGTTATTTCATCATG
CGATAGCTGCCCTGGTGGAATACAATAAAGTAGTAC

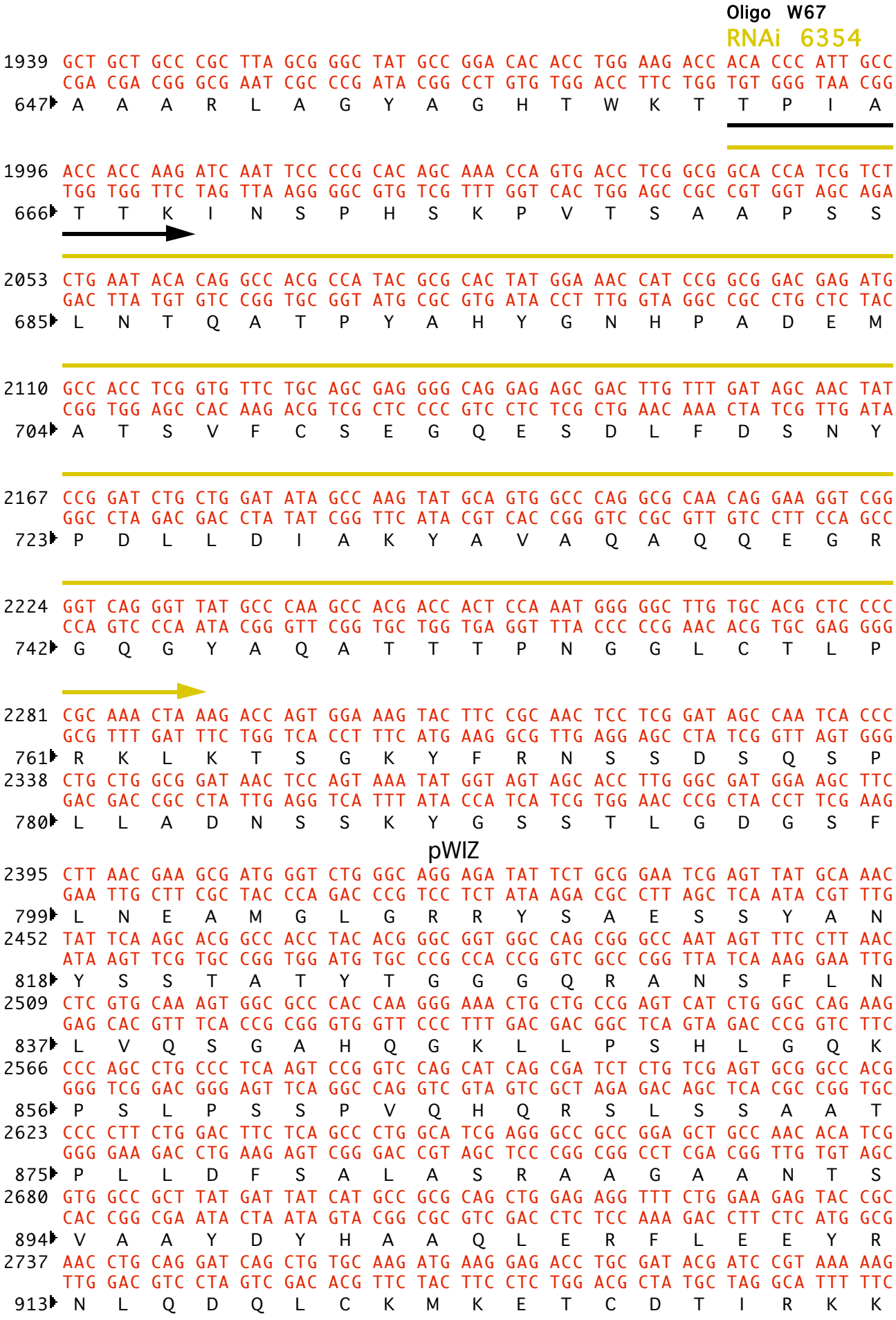
Kek3 ORF

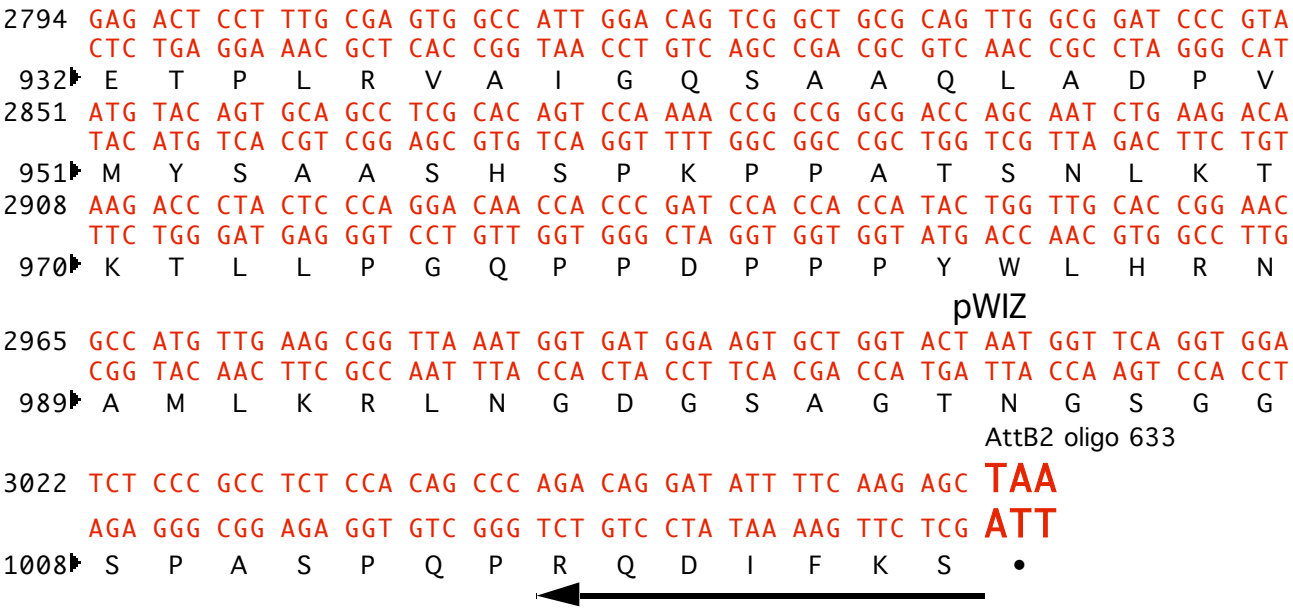
kek3/CG4192 ATG

AttB1 oligo W97

```
  1 ATG GCA GCG GGA AGA GCA GCC GCT ACG CTG GAG GCT CCG GGA CCG CCC AGC GGT CAG
    TAC CGT CGC CCT TCT CGT CGG CGA TGC GAC CTC CGA GGC CCT GGC GGG TCG CCA GTC
  1 M   A   A   G   R   A   A   A   T   L   E   A   P   G   P   P   S   G   Q

 58 GAC ATA GCC AGC GAC AAC AGC GCC CAG CGC CGC ACG CTG GCG ACG AAG GTG CGT CGA
    CTG TAT CGG TCG CTG TTG TCG CGG GTC GCG GCG TGC GAC CGC TGC TTC CAC GCA GCT
 20 D   I   A   S   D   N   S   A   Q   R   R   T   L   A   T   K   V   R   R
115 AAA GGG CCA CGC CCC CAA CGG CGC CTG CAC CCG CCC CTG CGC CCT CGC CTG CCG CTC
    TTT CCC GGT GCG GGG GTT GCC GCG GAC GTG GGC GGG GAC GCG GGA GCG GAC GGC GAG
 39 K   G   P   R   P   Q   R   R   L   H   P   P   L   R   P   R   L   P   L
172 CAT TTG CAC CTG CTA CTC TGG CTG CTG TGC TGC TGT TCT CAG CTG GGC CAG CTG AGG
    GTA AAC GTG GAC GAT GAG ACC GAC GAC ACG ACG ACA AGA GTC GAC CCG GTC GAC TCC
 58 H   L   H   L   L   L   W   L   L   C   C   C   S   Q   L   G   Q   L   R
229 GCC GAG TGT CCA GCG GTG TGC GAG TGC AAG TGG AAG AGT GGC AAG GAG TCC GTC TTG
    CGG CTC ACA GGT CGC CAC ACG CTC ACG TTC ACC TTC TCA CCG TTC CTC AGG CAG AAC
 77 A   E   C   P   A   V   C   E   C   K   W   K   S   G   K   E   S   V   L
                                                                LRR1
286 TGC CTT AAC GCC AAC CTA ACC CAC ATC CCG CAG CCG CTG GAC GCG GGA ACT CAG TTG
    ACG GAA TTG CGG TTG GAT TGG GTG TAG GGC GTC GGC GAC CTG CGC CCT TGA GTC AAC
 96 C   L   N   A   N   L   T   H   I   P   Q   P   L   D   A   G   T   Q   L
343 CTG GAC CTT AGC GGC AAT GAG ATC CAA CTA ATA CCC GAC GAT AGC TTC GCA ACC GCC
    GAC CTG GAA TCG CCG TTA CTC TAG GTT GAT TAT GGG CTG CTA TCG AAG CGT TGG CGG
115 L   D   L   S   G   N   E   I   Q   L   I   P   D   D   S   F   A   T   A
        LRR2
400 CAG TTG CTC AAC CTA CAG AAG GTG TAC CTG GCC AGG TGT CAC CTC CGG CTT ATC GAA
    GTC AAC GAG TTG GAT GTC TTC CAC ATG GAC CGG TCC ACA GTG GAG GCC GAA TAG CTT
134 Q   L   L   N   L   Q   K   V   Y   L   A   R   C   H   L   R   L   I   E
                        LRR3
457 CGC CAT GCC TTC CGT AAG CTG ATC AAT CTA GTG GAA CTG GAT CTA AGC CAG AAC CTG
    GCG GTA CGG AAG GCA TTC GAC TAG TTA GAT CAC CTT GAC CTA GAT TCG GTC TTG GAC
153 R   H   A   F   R   K   L   I   N   L   V   E   L   D   L   S   Q   N   L
                                                LRR4
514 CTC TCG GCA ATA CCC TCA TTG GCG CTC TAC CAT GTC TCA GAG CTA AGG GAG CTC CGA
    GAG AGC CGT TAT GGG AGT AAC CGC GAG ATG GTA CAG AGT CTC GAT TCC CTC GAG GCT
172 L   S   A   I   P   S   L   A   L   Y   H   V   S   E   L   R   E   L   R
                                                                LRR5
571 CTG AGT GGC AAT CCC ATA CTG AGA GTG CCA GAC GAT GCA TTT GGT CAT GTC CCA CAA
    GAC TCA CCG TTA GGG TAT GAC TCT CAC GGT CTG CTA CGT AAA CCA GTA CAG GGT GTT
191 L   S   G   N   P   I   L   R   V   P   D   D   A   F   G   H   V   P   Q
628 TTG GTG AAG CTG GAG CTG AGC GAC TGC CGC CTT TCG CAC ATC GCT GTG CGA GCA TTT
    AAC CAC TTC GAC CTC GAC TCG CTG ACG GCG GAA AGC GTG TAG CGA CAC GCT CGT AAA
210 L   V   K   L   E   L   S   D   C   R   L   S   H   I   A   V   R   A   F
            LRR6
685 GCC GGG CTG GAG AGC AGT CTG GAG TGG CTA AAA CTG GAT GGG AAT CGG CTG AGC GAG
    CGG CCC GAC CTC TCG TCA GAC CTC ACC GAT TTT GAC CTA CCC TTA GCC GAC TCG CTC
229 A   G   L   E   S   S   L   E   W   L   K   L   D   G   N   R   L   S   E
                                    LRR7
742 GTC AGG AGT GGC ACG ATC ACC TCG CTG GCT TCA CTG CAT GGT CTG GAG TTG GCG CGC
    CAG TCC TCA CCG TGC TAG TGG AGC GAC CGA AGT GAC GTA CCA GAC CTC AAC CGC GCG
248 V   R   S   G   T   I   T   S   L   A   S   L   H   G   L   E   L   A   R
        C Flank
799 AAT ACC TGG AAT TGC AGC TGC TCC TTG CGT CCT TTG AGG GCC TGG ATG CTG CAG CAG
    TTA TGG ACC TTA ACG TCG ACG AGG AAC GCA GGA AAC TCC CGG ACC TAC GAC GTC GTC
267 N   T   W   N   C   S   C   S   L   R   P   L   R   A   W   M   L   Q   Q
856 AAT ATA CCG AGT GGC ATA CCG CCA ACA TGT GAG TCT CCT CCT AGA TTG TCC GGG AGG
    TTA TAT GGC TCA CCG TAT GGC GGT TGT ACA CTC AGA GGA GGA TCT AAC AGG CCC TCC
286 N   I   P   S   G   I   P   P   T   C   E   S   P   P   R   L   S   G   R
```

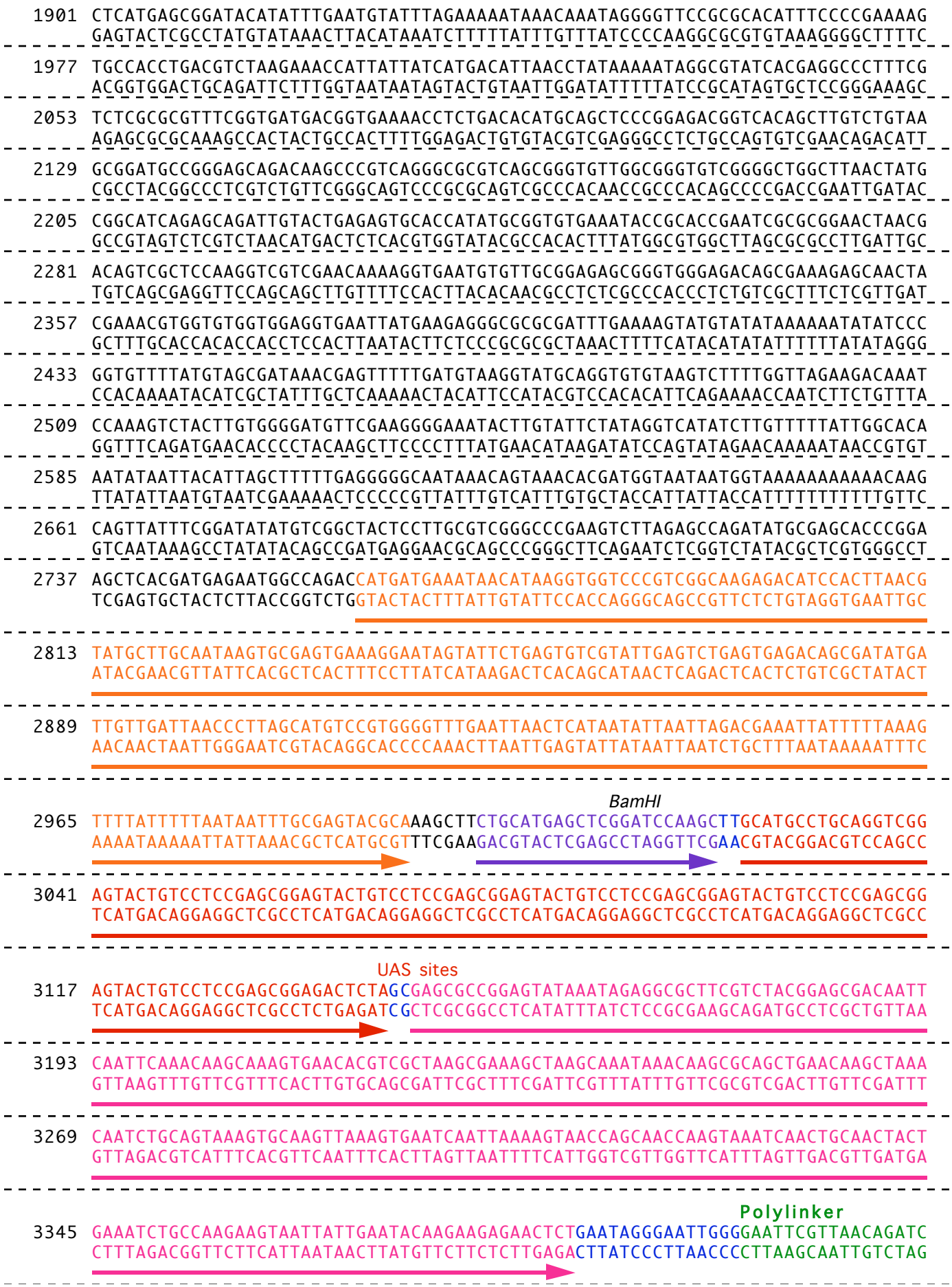

913 GCT TGG GAT AAG CTC GAT GTA GAT GAC TTT GCG TGC GTT CCA CAA ATT GTG GCC ACG
CGA ACC CTA TTC GAG CTA CAT CTA CTG AAA CGC ACG CAA GGT GTT TAA CAC CGG TGC
305 A W D K L D V D D F A C V P Q I V A T
970 GAC ACC ACA GCG CAT GGA GTG GAG GGC AGG AAC ATA ACC ATG AGC TGC TAC GTG GAA
CTG TGG TGT CGC GTA CCT CAC CTC CCG TCC TTG TAT TGG TAC TCG ACG ATG CAC CTT
324 D T T A H G V E G R N I T M S C Y V E
1027 GGA GTA CCC CAA CCG GCT GTC AAG TGG CTG CTT AAA AAC CGA CTG ATA GCC AAT CTC
CCT CAT GGG GTT GGC CGA CAG TTC ACC GAC GAA TTT TTG GCT GAC TAT CGG TTA GAG
343 G V P Q P A V K W L L K N R L I A N L
1084 AGT GCT GGC GGG GAT GGT GAC TCC GAT TCG GAG CCC AGG ACA GCG GCA GCA ACT CAG
TCA CGA CCG CCC CTA CCA CTG AGG CTA AGC CTC GGG TCC TGT CGC CGT CGT TGA GTC
362 S A G G D G D S D S E P R T A A A T Q
1141 GGT AGG AAG ACC TAT GTG GTC AAC ATG CTG AGA AAT GCC TCG AAC CTG ACC ATT CTC
CCA TCC TTC TGG ATA CAC CAG TTG TAC GAC TCT TTA CGG AGC TTG GAC TGG TAA GAG
381 G R K T Y V V N M L R N A S N L T I L
1198 ACG GCT GAC ATG CAG GAT GCC GGG ATC TAT ACG TGT GCG GCG GAA AAT AAG GCT GGA
TGC CGA CTG TAC GTC CTA CGG CCC TAG ATA TGC ACA CGC CGC CTT TTA TTC CGA CCT
400 T A D M Q D A G I Y T C A A E N K A G
1255 AAA GTG GAG GCC AGT GTG ACT CTG GCA GTA TCC CGT AGA CCC CCG GAA GCT CCG TGG
TTT CAC CTC CGG TCA CAC TGA GAC CGT CAT AGG GCA TCT GGG GGC CTT CGA GGC ACC
419 K V E A S V T L A V S R R P P E A P W
Tm?
1312 GGC GTA AGA ATT ATT CTG CTG GGG GCG GTA GCC GCT CTG CTC CTC GTC GGT GGA TCC
CCG CAT TCT TAA TAA GAC GAC CCC CGC CAT CGG CGA GAC GAG GAG CAG CCA CCT AGG
438 G V R I I L L G A V A A L L L V G G S
IC?
1369 TCC TTT GCG GCC ATT TGC TTG TGT TCC CTA CAA AGG CGA AGA AAG CTG CGT CTC TGG
AGG AAA CGC CGG TAA ACG AAC ACA AGG GAT GTT TCC GCT TCT TTC GAC GCA GAG ACC
457 S F A A I C L C S L Q R R R K L R L W
1426 AAC TCT GTA CCT CCT GTG AGG AGA AGC GAG AGC TAC GAA AAG ATC GAG ATG ACG GCC
TTG AGA CAT GGA GGA CAC TCC TCT TCG CTC TCG ATG CTT TTC TAG CTC TAC TGC CGG
476 N S V P P V R R S E S Y E K I E M T A
1483 AGA ACG AGA CCG GAT CTG GGA GGA GGG GCT AGT TGC GGA GGC GGC AGT GCC ACG GGC
TCT TGC TCT GGC CTA GAC CCT CCT CCC CGA TCA ACG CCT CCG CCG TCA CGG TGC CCG
495 R T R P D L G G G A S C G G G S A T G
1540 GCC GGA CTC TTT CAC GAT GCC GAG GAG CAG GGC TAT CTG CGG GCA GCT CAT ACG CCA
CGG CCT GAG AAA GTG CTA CGG CTC CTC GTC CCG ATA GAC GCC CGT CGA GTA TGC GGT
514 A G L F H D A E E Q G Y L R A A H T P
1597 CTA AAT GAC AAC GAT GCC GGG CAG GCG GCG GCC ATC GTA AAT CCG AGT GCA GGA AGT
GAT TTA CTG TTG CTA CGG CCC GTC CGC CGC CGG TAG CAT TTA GGC TCA CGT CCT TCA
533 L N D N D A G Q A A A I V N P S A G S
1654 GCA CAG CGA AGA AAT GGA GAC TAC CTG CAC GTG TCC ACC CAC TGC GAT GAT GAG GAG
CGT GTC GCT TCT TTA CCT CTG ATG GAC GTG CAC AGG TGG GTG ACG CTA CTA CTC CTC
552 A Q R R N G D Y L H V S T H C D D E E
1711 GAG GAC CAA CAG CTG CAT CAC CAC CCA CAA CAG CAG CCC GCG AGC CAG CAC CAC CCA
CTC CTG GTT GTC GAC GTA GTG GTG GGT GTT GTC GTC GGG CGC TCG GTC GTG GTG GGT
571 E D Q Q L H H H P Q Q Q P A S Q H H P
1768 CAT CCC AAT CAG CAG CAG CAT CAG CAA AGG AAG GGC TCC CAG GGC CAT GTT GTC TCC
GTA GGG TTA GTC GTC GTC GTA GTC GTT TCC TTC CCG AGG GTC CCG GTA CAA CAG AGG
590 H P N Q Q Q H Q Q R K G S Q G H V V S
1825 GCA TCC GGG GCG AAT AAT TCA GCA CCG CTG GAG GAA ACG GAT CTG CAC ATA CCG CGC
CGT AGG CCC CGC TTA TTA AGT CGT GGC GAC CTC CTT TGC CTA GAC GTG TAT GGC GCG
609 A S G A N N S A P L E E T D L H I P R
splice
1882 CTC ATC GAC ATC GGC GGC ACC GAT TCC GCA TCG AGT TCA ATC TCC AGC CAG GTG GAC
GAG TAG CTG TAG CCG CCG TGG CTA AGG CGT AGC TCA AGT TAG AGG TCG GTC CAC CTG
628 L I D I G G T D S A S S S I S S Q V D

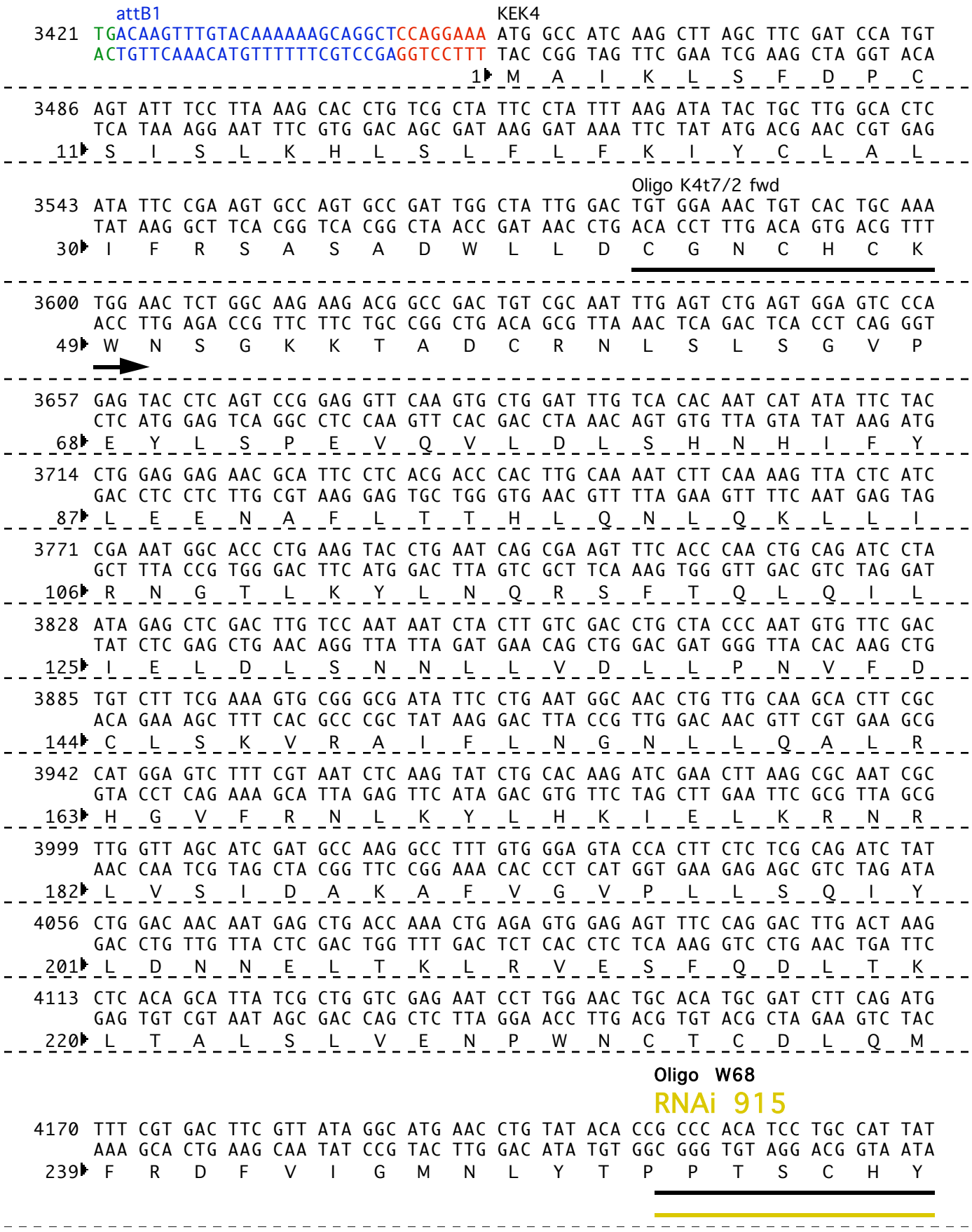

Oligo W67
RNAi 6354
1939 GCT GCT GCC CGC TTA GCG GGC TAT GCC GGA CAC ACC TGG AAG ACC ACA CCC ATT GCC
CGA CGA CGG GCG AAT CGC CCG ATA CGG CCT GTG TGG ACC TTC TGG TGT GGG TAA CGG
647 A A A R L A G Y A G H T W K T T P I A
1996 ACC ACC AAG ATC AAT TCC CCG CAC AGC AAA CCA GTG ACC TCG GCG GCA CCA TCG TCT
TGG TGG TTC TAG TTA AGG GGC GTG TCG TTT GGT CAC TGG AGC CGC CGT GGT AGC AGA
666 T T K I N S P H S K P V T S A A P S S
2053 CTG AAT ACA CAG GCC ACG CCA TAC GCG CAC TAT GGA AAC CAT CCG GCG GAC GAG ATG
GAC TTA TGT GTC CGG TGC GGT ATG CGC GTG ATA CCT TTG GTA GGC CGC CTG CTC TAC
685 L N T Q A T P Y A H Y G N H P A D E M
2110 GCC ACC TCG GTG TTC TGC AGC GAG GGG CAG GAG AGC GAC TTG TTT GAT AGC AAC TAT
CGG TGG AGC CAC AAG ACG TCG CTC CCC GTC CTC TCG CTG AAC AAA CTA TCG TTG ATA
704 A T S V F C S E G Q E S D L F D S N Y
2167 CCG GAT CTG CTG GAT ATA GCC AAG TAT GCA GTG GCC CAG GCG CAA CAG GAA GGT CGG
GGC CTA GAC GAC CTA TAT CGG TTC ATA CGT CAC CGG GTC CGC GTT GTC CTT CCA GCC
723 P D L L D I A K Y A V A Q A Q Q E G R
2224 GGT CAG GGT TAT GCC CAA GCC ACG ACC ACT CCA AAT GGG GGC TTG TGC ACG CTC CCC
CCA GTC CCA ATA CGG GTT CGG TGC TGG TGA GGT TTA CCC CCG AAC ACG TGC GAG GGG
742 G Q G Y A Q A T T T P N G G L C T L P
2281 CGC AAA CTA AAG ACC AGT GGA AAG TAC TTC CGC AAC TCC TCG GAT AGC CAA TCA CCC
GCG TTT GAT TTC TGG TCA CCT TTC ATG AAG GCG TTG AGG AGC CTA TCG GTT AGT GGG
761 R K L K T S G K Y F R N S S D S Q S P
2338 CTG CTG GCG GAT AAC TCC AGT AAA TAT GGT AGT AGC ACC TTG GGC GAT GGA AGC TTC
GAC GAC CGC CTA TTG AGG TCA TTT ATA CCA TCA TCG TGG AAC CCG CTA CCT TCG AAG
780 L L A D N S S K Y G S S T L G D G S F
pWIZ
2395 CTT AAC GAA GCG ATG GGT CTG GGC AGG AGA TAT TCT GCG GAA TCG AGT TAT GCA AAC
GAA TTG CTT CGC TAC CCA GAC CCG TCC TCT ATA AGA CGC CTT AGC TCA ATA CGT TTG
799 L N E A M G L G R R Y S A E S S Y A N
2452 TAT TCA AGC ACG GCC ACC TAC ACG GGC GGT GGC CAG CGG GCC AAT AGT TTC CTT AAC
ATA AGT TCG TGC CGG TGG ATG TGC CCG CCA CCG GTC GCC CGG TTA TCA AAG GAA TTG
818 Y S S T A T Y T G G G Q R A N S F L N
2509 CTC GTG CAA AGT GGC GCC CAC CAA GGG AAA CTG CTG CCG AGT CAT CTG GGC CAG AAG
GAG CAC GTT TCA CCG CGG GTG GTT CCC TTT GAC GAC GGC TCA GTA GAC CCG GTC TTC
837 L V Q S G A H Q G K L L P S H L G Q K
2566 CCC AGC CTG CCC TCA AGT CCG GTC CAG CAT CAG CGA TCT CTG TCG AGT GCG GCC ACG
GGG TCG GAC GGG AGT TCA GGC CAG GTC GTA GTC GCT AGA GAC AGC TCA CGC CGG TGC
856 P S L P S S P V Q H Q R S L S S A A T
2623 CCC CTT CTG GAC TTC TCA GCC CTG GCA TCG AGG GCC GCC GGA GCT GCC AAC ACA TCG
GGG GAA GAC CTG AAG AGT CGG GAC CGT AGC TCC CGG CGG CCT CGA CGG TTG TGT AGC
875 P L L D F S A L A S R A A G A A N T S
2680 GTG GCC GCT TAT GAT TAT CAT GCC GCG CAG CTG GAG AGG TTT CTG GAA GAG TAC CGC
CAC CGG CGA ATA CTA ATA GTA CGG CGC GTC GAC CTC TCC AAA GAC CTT CTC ATG GCG
894 V A A Y D Y H A A Q L E R F L E E Y R
2737 AAC CTG CAG GAT CAG CTG TGC AAG ATG AAG GAG ACC TGC GAT ACG ATC CGT AAA AAG
TTG GAC GTC CTA GTC GAC ACG TTC TAC TTC CTC TGG ACG CTA TGC TAG GCA TTT TTC
913 N L Q D Q L C K M K E T C D T I R K K

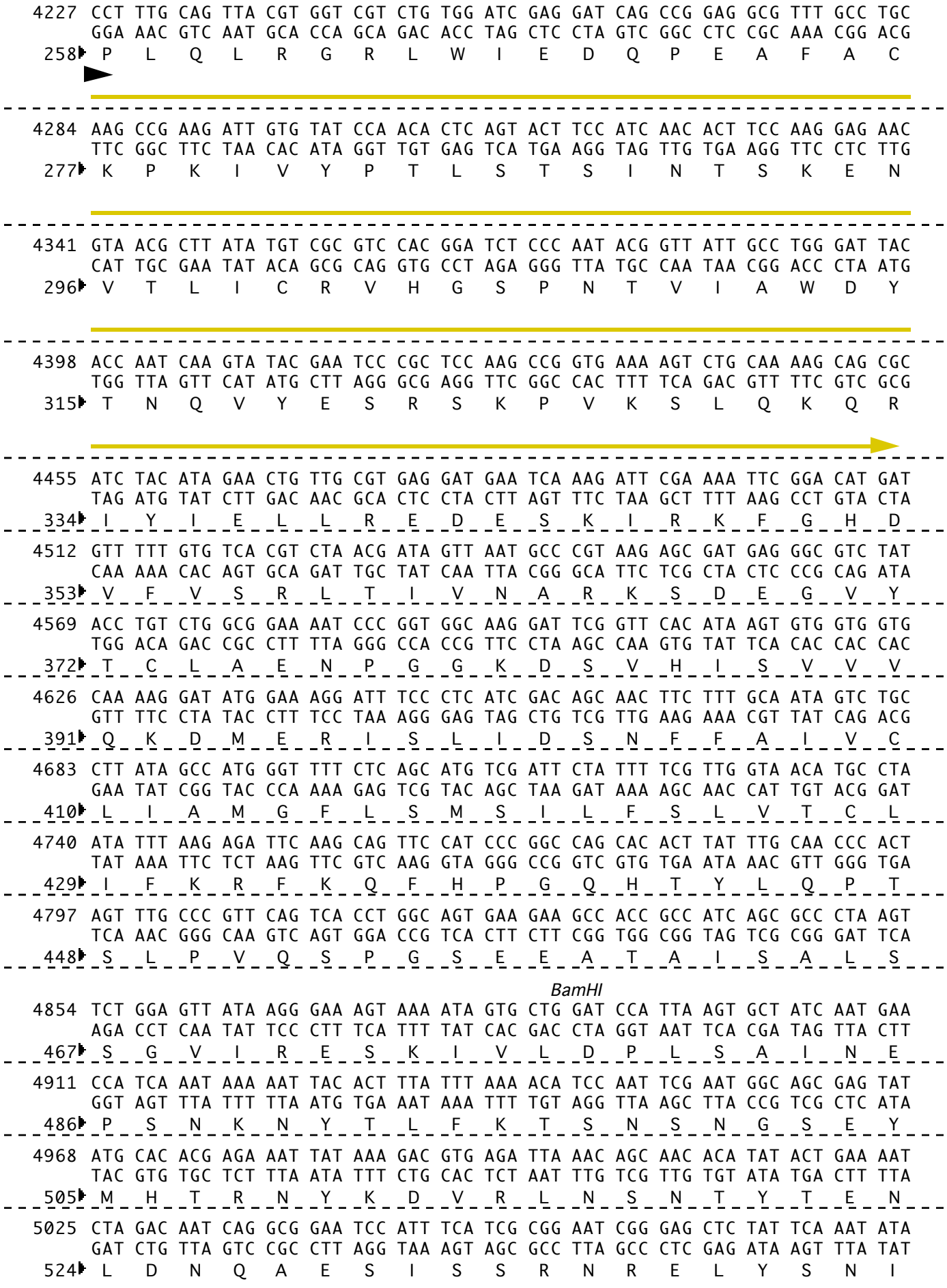

2794 GAG ACT CCT TTG CGA GTG GCC ATT GGA CAG TCG GCT GCG CAG TTG GCG GAT CCC GTA
CTC TGA GGA AAC GCT CAC CGG TAA CCT GTC AGC CGA CGC GTC AAC CGC CTA GGG CAT
932 E T P L R V A I G Q S A A Q L A D P V
2851 ATG TAC AGT GCA GCC TCG CAC AGT CCA AAA CCG CCG GCG ACC AGC AAT CTG AAG ACA
TAC ATG TCA CGT CGG AGC GTG TCA GGT TTT GGC GGC CGC TGG TCG TTA GAC TTC TGT
951 M Y S A A S H S P K P P A T S N L K T
2908 AAG ACC CTA CTC CCA GGA CAA CCA CCC GAT CCA CCA CCA TAC TGG TTG CAC CGG AAC
TTC TGG GAT GAG GGT CCT GTT GGT GGG CTA GGT GGT GGT ATG ACC AAC GTG GCC TTG
970 K T L L P G Q P P D P P P Y W L H R N
pWIZ
2965 GCC ATG TTG AAG CGG TTA AAT GGT GAT GGA AGT GCT GGT ACT AAT GGT TCA GGT GGA
CGG TAC AAC TTC GCC AAT TTA CCA CTA CCT TCA CGA CCA TGA TTA CCA AGT CCA CCT
989 A M L K R L N G D G S A G T N G S G G
AttB2 oligo 633
3022 TCT CCC GCC TCT CCA CAG CCC AGA CAG GAT ATT TTC AAG AGC TAA
AGA GGG CGG AGA GGT GTC GGG TCT GTC CTA TAA AAG TTC TCG ATT
1008 S P A S P Q P R Q D I F K S •

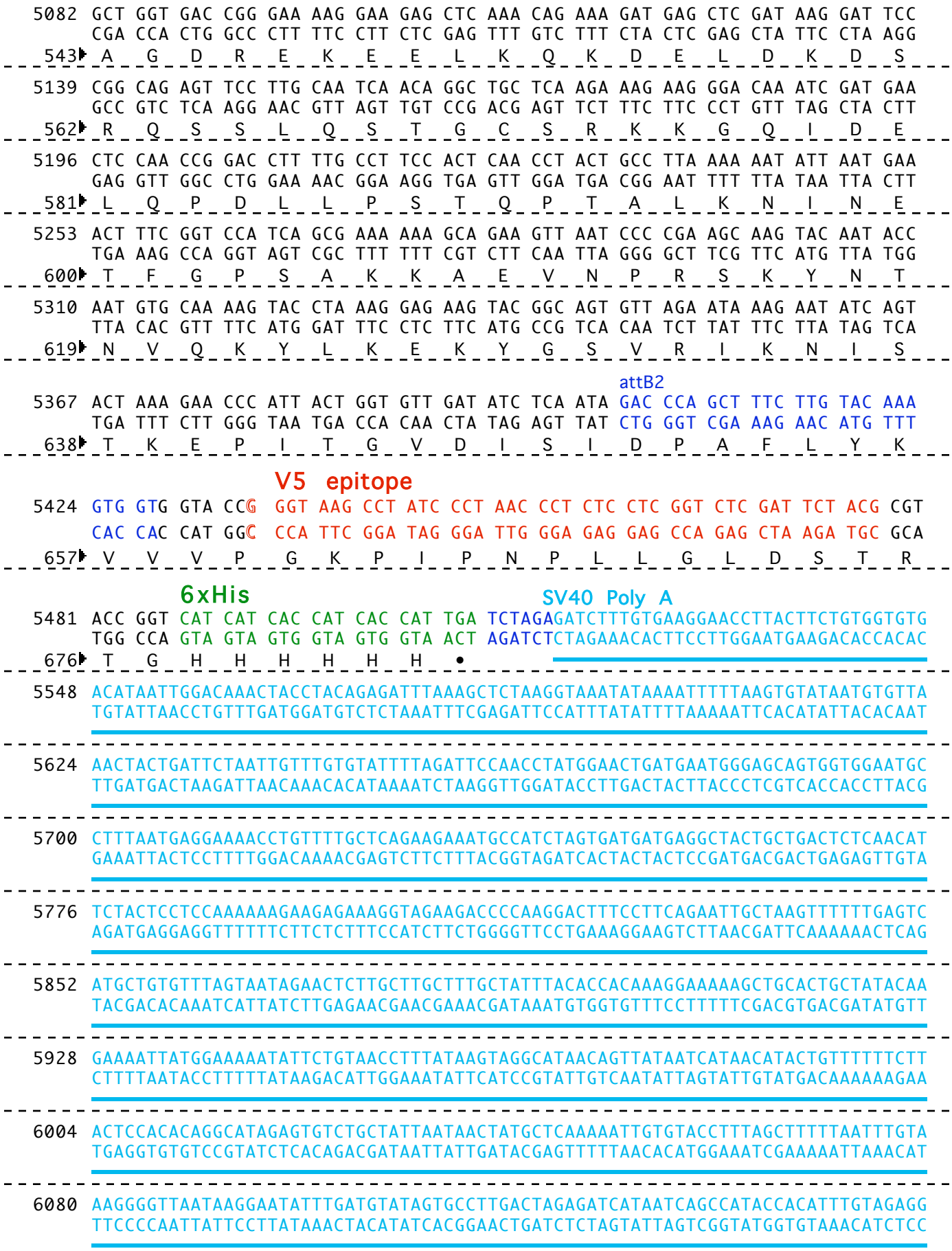
pUAS Kek4

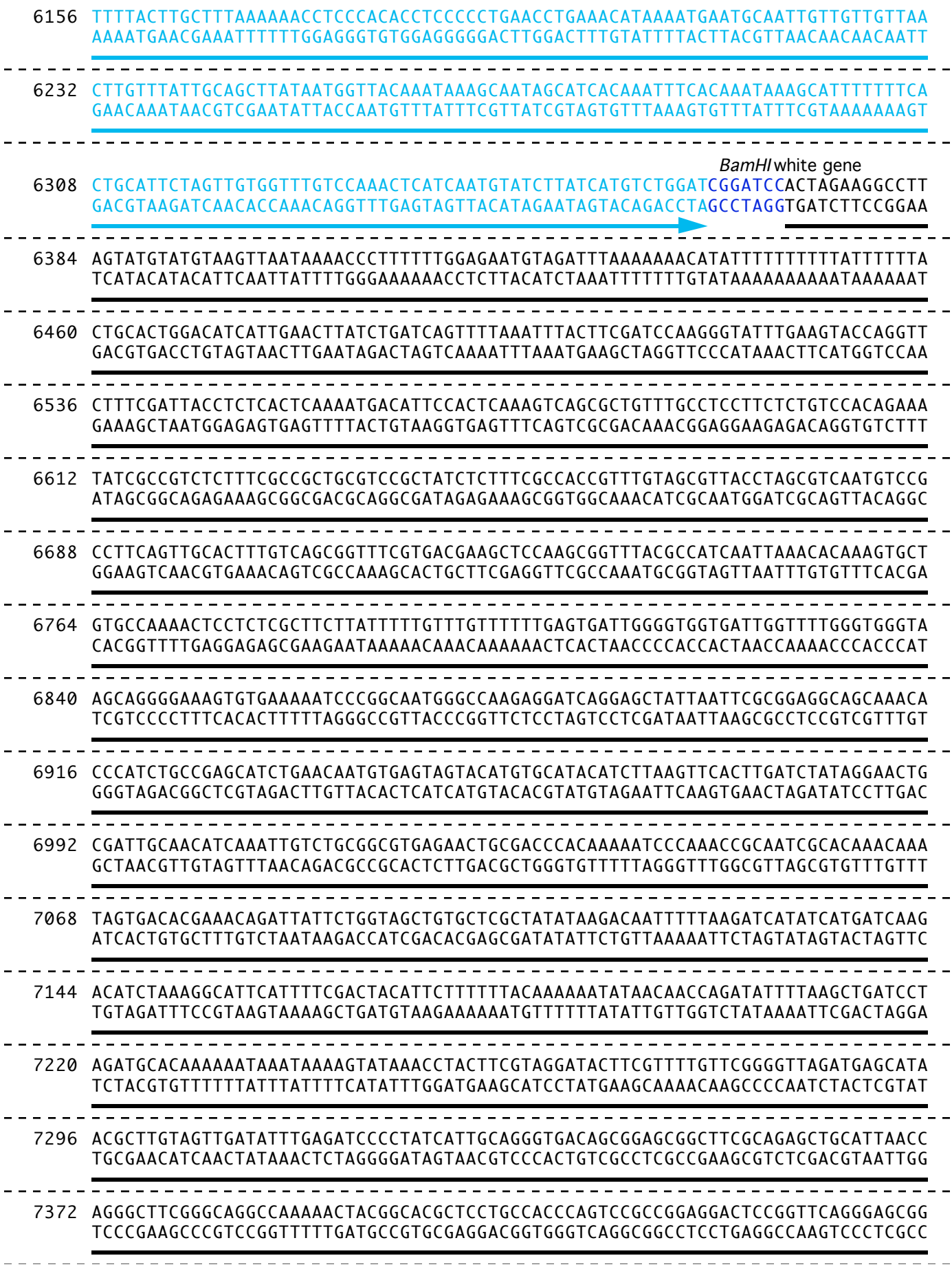
```
     1 GGCCAGACCCACGTAGTCCAGCGGCAGATCGGCGGCGGAGAAGTTAAGCGTCTCCAGGATGACCTTGCCCGAACTG
       CCGGTCTGGGTGCATCAGGTCGCCGTCTAGCCGCCGCCTCTTCAATTCGCAGAGGTCCTACTGGAACGGGCTTGAC

    77 GGGCACGTGGTGTTCGACGATGTGCAGCTAATTTCGCCCGGCTCCACGTCCGCCCATTGGTTAATCAGCAGACCCT
       CCCGTGCACCACAAGCTGCTACACGTCGATTAAAGCGGGCCGAGGTGCAGGCGGGTAACCAATTAGTCGTCTGGGA

   153 CGTTGGCGTAACGGAACCATGAGAGGTACGACAACCATTTGAGGTATACTGGCACCGAGCCCGAGTTCAAGAAGAA
       GCAACCGCATTGCCTTGGTACTCTCCATGCTGTTGGTAAACTCCATATGACCGTGGCTCGGGCTCAAGTTCTTCTT

   229 GGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCG
       CCGCAAAAAGGTATCCGAGGCGGGGGGACTGCTCGTAGTGTTTTTAGCTGCGAGTTCAGTCTCCACCGCTTTGGGC

   305 ACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTA
       TGTCCTGATATTTCTATGGTCCGCAAAGGGGGACCTTCGAGGGAGCACGCGAGAGGACAAGGCTGGGACGGCGAAT

   381 CCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTC
       GGCCTATGGACAGGCGGAAAGAGGGAAGCCCTTCGCACCGCGAAAGAGTTACGAGTGCGACATCCATAGAGTCAAG

   457 GGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGT
       CCACATCCAGCAAGCGAGGTTCGACCCGACACACGTGCTTGGGGGGCAAGTCGGGCTGGCGACGCGGAATAGGCCA

   533 AACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA
       TTGATAGCAGAACTCAGGTTGGGCCATTCTGTGCTGAATAGCGGTGACCGTCGTCGGTGACCATTGTCCTAATCGT

   609 GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATT
       CTCGCTCCATACATCCGCCACGATGTCTCAAGAACTTCACCACCGGATTGATGCCGATGTGATCTTCCTGTCATAA

   685 TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
       ACCATAGACGCGAGACGACTTCGGTCAATGGAAGCCTTTTTCTCAACCATCGAGAACTAGGCCGTTTGTTTGGTGG

   761 GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGA
       CGACCATCGCCACCAAAAAAACAAACGTTCGTCGTCTAATGCGCGTCTTTTTTTCCTAGAGTTCTTCTAGGAAACT

   837 TCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAG
       AGAAAAGATGCCCCAGACTGCGAGTCACCTTGCTTTTGAGTGCAATTCCCTAAAACCAGTACTCTAATAGTTTTTC

   913 GATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCT
       CTAGAAGTGGATCTAGGAAAATTTAATTTTTACTTCAAAATTTAGTTAGATTTCATATATACTCATTTGAACCAGA

   989 GACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGAC
       CTGTCAATGGTTACGAATTAGTCACTCCGTGGATAGAGTCGCTAGACAGATAAAGCAAGTAGGTATCAACGGACTG

  1065 TCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC
       AGGGGCAGCACATCTATTGATGCTATGCCCTCCCGAATGGTAGACCGGGGTCACGACGTTACTATGGCGCTCTGGG

  1141 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACT
       TGCGAGTGGCCGAGGTCTAAATAGTCGTTATTTGGTCGGTCGGCCTTCCCGGCTCGCGTCTTCACCAGGACGTTGA

  1217 TTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCA
       AATAGGCGGAGGTAGGTCAGATAATTAACAACGGCCCTTCGATCTCATTCATCAAGCGGTCAATTATCAAACGCGT

  1293 ACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCA
       TGCAACAACGGTAACGATGTCCGTAGCACCACAGTGCGAGCAGCAAACCATACCGAAGTAAGTCGAGGCCAAGGGT

  1369 ACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTC
       TGCTAGTTCCGCTCAATGTACTAGGGGGTACAACACGTTTTTTCGCCAATCGAGGAAGCCAGGAGGCTAGCAACAG

  1445 AGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCG
       TCTTCATTCAACCGGCGTCACAATAGTGAGTACCAATACCGTCGTGACGTATTAAGAGAATGACAGTACGGTAGGC

  1521 TAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTC
       ATTCTACGAAAAGACACTGACCACTCATGAGTTGGTTCAGTAAGACTCTTATCACATACGCCGCTGGCTCAACGAG

  1597 TTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCT
       AACGGGCCGCAGTTGTGCCCTATTATGGCGCGGTGTATCGTCTTGAAATTTTCACGAGTAGTAACCTTTTGCAAGA

  1673 TCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGAT
       AGCCCCGCTTTTGAGAGTTCCTAGAATGGCGACAACTCTAGGTCAAGCTACATTGGGTGAGCACGTGGGTTGACTA

  1749 CTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAAT
       GAAGTCGTAGAAAATGAAAGTGGTCGCAAAGACCCACTCGTTTTTGTCCTTCCGTTTTACGGCGTTTTTTCCCTTA

  1825 AAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGT
       TTCCCGCTGTGCCTTTACAACTTATGAGTATGAGAAGGAAAAAGTTATAATAACTTCGTAAATAGTCCCAATAACA
```

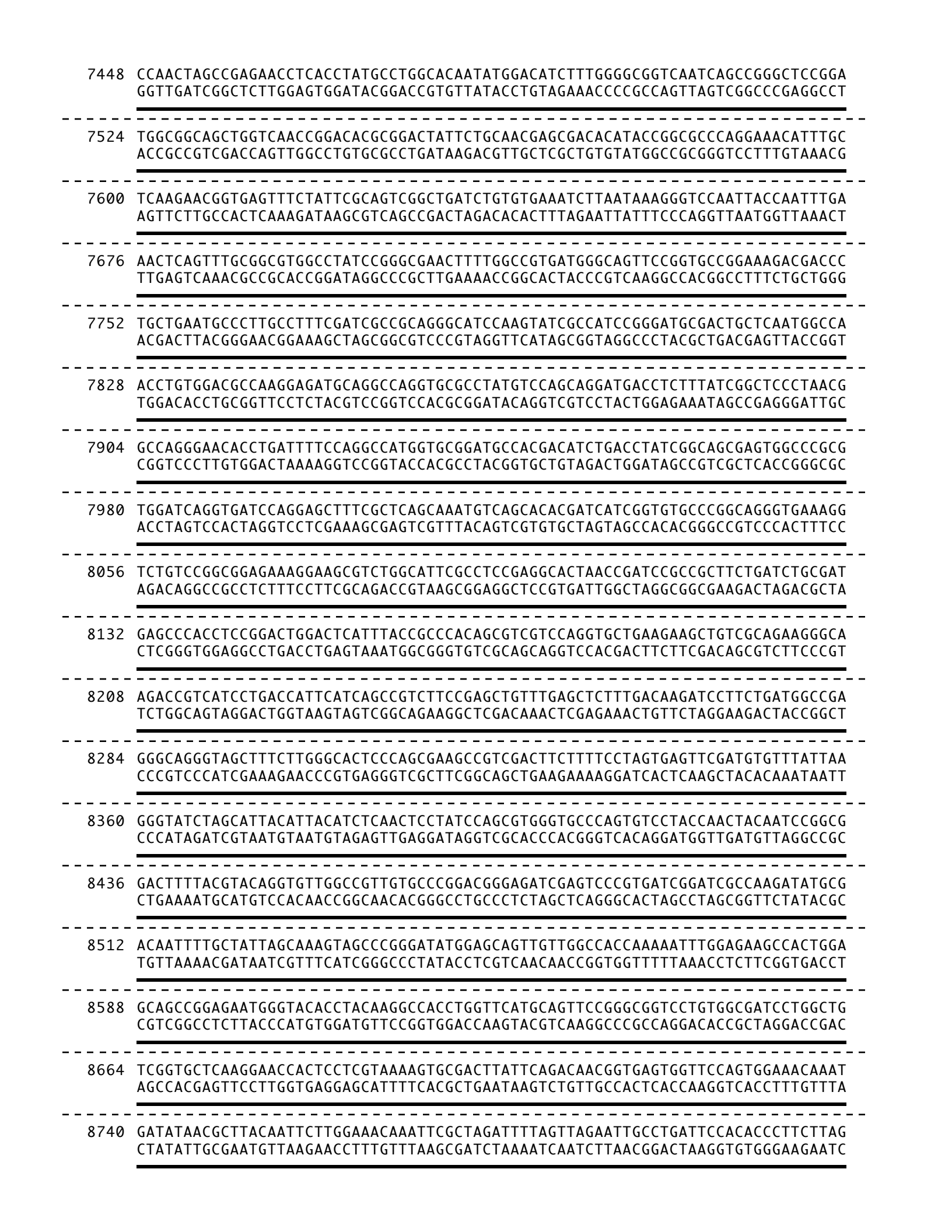

1901 CTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAG
GAGTACTCGCCTATGTATAAACTTACATAAATCTTTTTATTTGTTTATCCCCAAGGCGCGTGTAAAGGGGCTTTTC
1977 TGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCG
ACGGTGGACTGCAGATTCTTTGGTAATAATAGTACTGTAATTGGATATTTTTATCCGCATAGTGCTCCGGGAAAGC
2053 TCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAA
AGAGCGCGCAAAGCCACTACTGCCACTTTTGGAGACTGTGTACGTCGAGGGCCTCTGCCAGTGTCGAACAGACATT
2129 GCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATG
CGCCTACGGCCCTCGTCTGTTCGGGCAGTCCCGCGCAGTCGCCCACAACCGCCCACAGCCCCGACCGAATTGATAC
2205 CGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACCGAATCGCGCGGAACTAACG
GCCGTAGTCTCGTCTAACATGACTCTCACGTGGTATACGCCACACTTTATGGCGTGGCTTAGCGCGCCTTGATTGC
2281 ACAGTCGCTCCAAGGTCGTCGAACAAAAGGTGAATGTGTTGCGGAGAGCGGGTGGGAGACAGCGAAAGAGCAACTA
TGTCAGCGAGGTTCCAGCAGCTTGTTTTCCACTTACACAACGCCTCTCGCCCACCCTCTGTCGCTTTCTCGTTGAT
2357 CGAAACGTGGTGTGGTGGAGGTGAATTATGAAGAGGGCGCGCGATTTGAAAAGTATGTATATAAAAAATATATCCC
GCTTTGCACCACACCACCTCCACTTAATACTTCTCCCGCGCGCTAAACTTTTCATACATATATTTTTTATATAGGG
2433 GGTGTTTTATGTAGCGATAAACGAGTTTTTGATGTAAGGTATGCAGGTGTGTAAGTCTTTTGGTTAGAAGACAAAT
CCACAAAATACATCGCTATTTGCTCAAAAACTACATTCCATACGTCCACACATTCAGAAAACCAATCTTCTGTTTA
2509 CCAAAGTCTACTTGTGGGGATGTTCGAAGGGGAAATACTTGTATTCTATAGGTCATATCTTGTTTTTATTGGCACA
GGTTTCAGATGAACACCCCTACAAGCTTCCCCTTTATGAACATAAGATATCCAGTATAGAACAAAAATAACCGTGT
2585 AATATAATTACATTAGCTTTTTGAGGGGGCAATAAACAGTAAACACGATGGTAATAATGGTAAAAAAAAAAACAAG
TTATATTAATGTAATCGAAAAACTCCCCCGTTATTTGTCATTTGTGCTACCATTATTACCATTTTTTTTTTTGTTC
2661 CAGTTATTTCGGATATATGTCGGCTACTCCTTGCGTCGGGCCCGAAGTCTTAGAGCCAGATATGCGAGCACCCGGA
GTCAATAAAGCCTATATACAGCCGATGAGGAACGCAGCCCGGGCTTCAGAATCTCGGTCTATACGCTCGTGGGCCT
2737 AGCTCACGATGAGAATGGCCAGACCATGATGAAATAACATAAGGTGGTCCCGTCGGCAAGAGACATCCACTTAACG
TCGAGTGCTACTCTTACCGGTCTGGTACTACTTTATTGTATTCCACCAGGGCAGCCGTTCTCTGTAGGTGAATTGC
2813 TATGCTTGCAATAAGTGCGAGTGAAAGGAATAGTATTCTGAGTGTCGTATTGAGTCTGAGTGAGACAGCGATATGA
ATACGAACGTTATTCACGCTCACTTTCCTTATCATAAGACTCACAGCATAACTCAGACTCACTCTGTCGCTATACT
2889 TTGTTGATTAACCCTTAGCATGTCCGTGGGGTTTGAATTAACTCATAATATTAATTAGACGAAATTATTTTTAAAG
AACAACTAATTGGGAATCGTACAGGCACCCCAAACTTAATTGAGTATTATAATTAATCTGCTTTAATAAAAATTTC
BamHI
2965 TTTTATTTTTAATAATTTGCGAGTACGCAAAGCTTCTGCATGAGCTCGGATCCAAGCTTGCATGCCTGCAGGTCGG
AAAATAAAAATTATTAAACGCTCATGCGTTTCGAAGACGTACTCGAGCCTAGGTTCGAACGTACGGACGTCCAGCC
3041 AGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAGCGG
TCATGACAGGAGGCTCGCCTCATGACAGGAGGCTCGCCTCATGACAGGAGGCTCGCCTCATGACAGGAGGCTCGCC
UAS sites
3117 AGTACTGTCCTCCGAGCGGAGACTCTAGCGAGCGCCGGAGTATAAATAGAGGCGCTTCGTCTACGGAGCGACAATT
TCATGACAGGAGGCTCGCCTCTGAGATCGCTCGCGGCCTCATATTTATCTCCGCGAAGCAGATGCCTCGCTGTTAA
3193 CAATTCAAACAAGCAAAGTGAACACGTCGCTAAGCGAAAGCTAAGCAAATAAACAAGCGCAGCTGAACAAGCTAAA
GTTAAGTTTGTTCGTTTCACTTGTGCAGCGATTCGCTTTCGATTCGTTTATTTGTTCGCGTCGACTTGTTCGATTT
3269 CAATCTGCAGTAAAGTGCAAGTTAAAGTGAATCAATTAAAAGTAACCAGCAACCAAGTAAATCAACTGCAACTACT
GTTAGACGTCATTTCACGTTCAATTTCACTTAGTTAATTTTCATTGGTCGTTGGTTCATTTAGTTGACGTTGATGA
Polylinker
3345 GAAATCTGCCAAGAAGTAATTATTGAATACAAGAAGAGAACTCTGAATAGGGAATTGGGGAATTCGTTAACAGATC
CTTTAGACGGTTCTTCATTAATAACTTATGTTCTTCTCTTGAGACTTATCCCTTAACCCCTTAAGCAATTGTCTAG

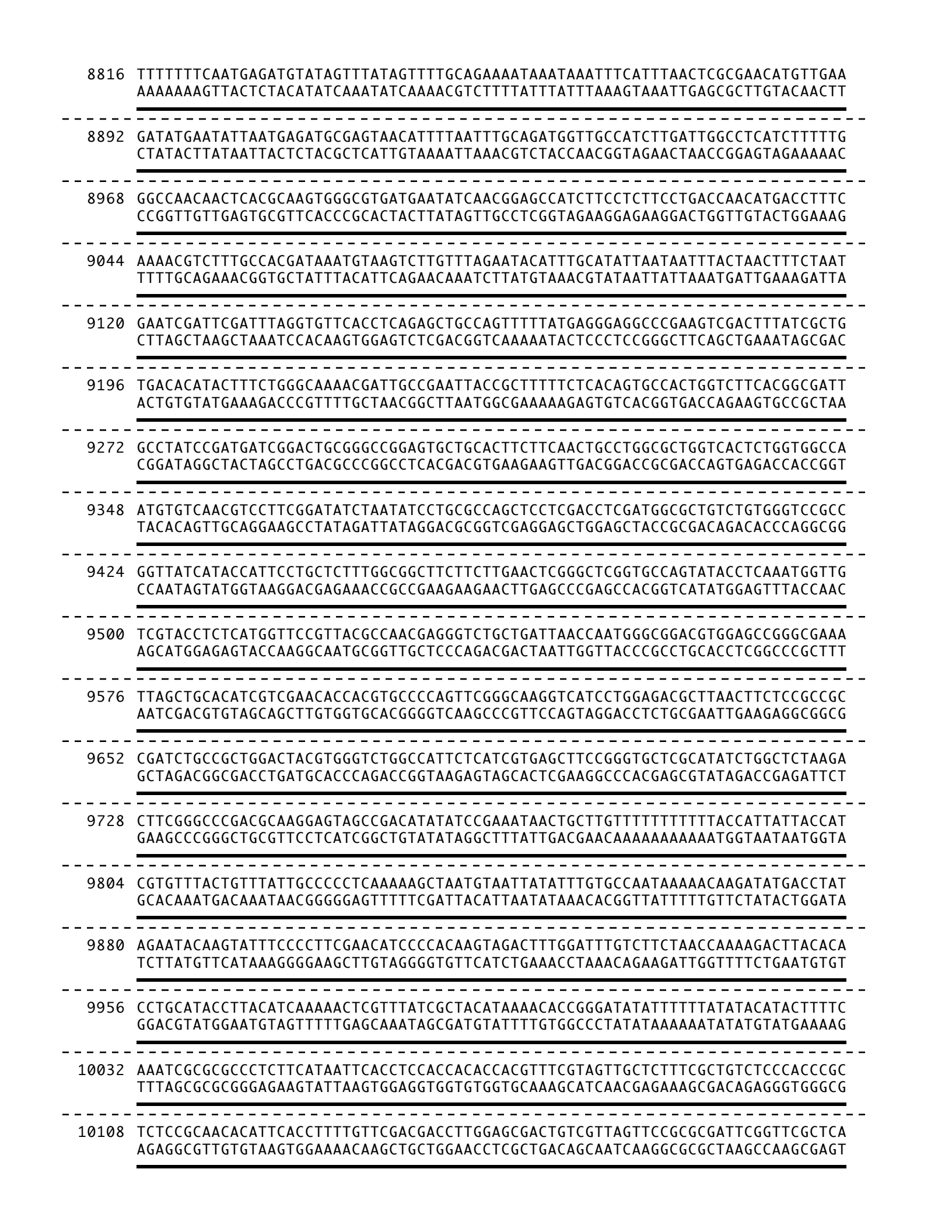

attB1
KEK4
Oligo K4t7/2 fwd
Oligo W68
RNAi 915

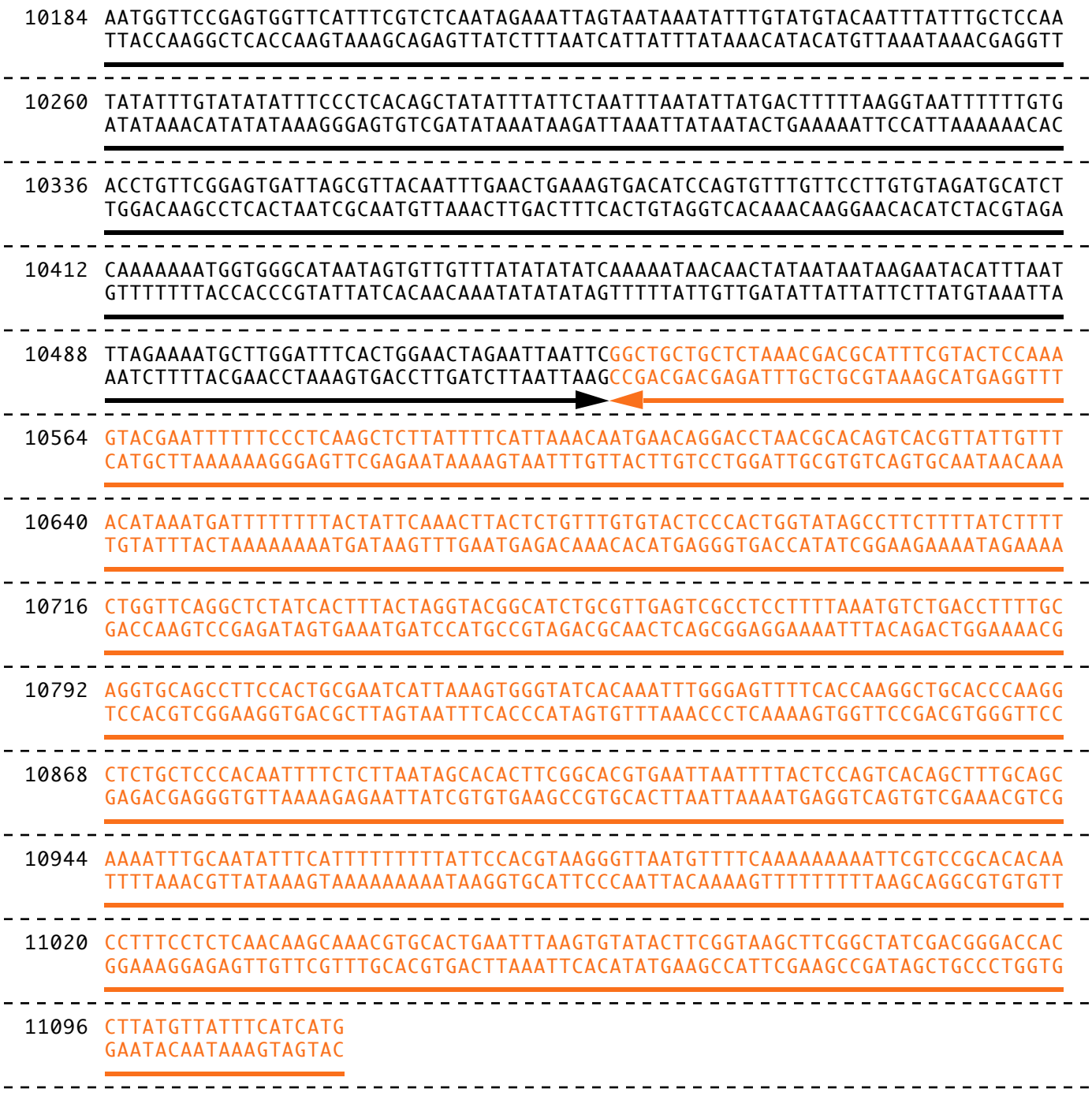

4227 CCT TTG CAG TTA CGT GGT CGT CTG TGG ATC GAG GAT CAG CCG GAG GCG TTT GCC TGC
GGA AAC GTC AAT GCA CCA GCA GAC ACC TAG CTC CTA GTC GGC CTC CGC AAA CGG ACG
258 P L Q L R G R L W I E D Q P E A F A C
4284 AAG CCG AAG ATT GTG TAT CCA ACA CTC AGT ACT TCC ATC AAC ACT TCC AAG GAG AAC
TTC GGC TTC TAA CAC ATA GGT TGT GAG TCA TGA AGG TAG TTG TGA AGG TTC CTC TTG
277 K P K I V Y P T L S T S I N T S K E N
4341 GTA ACG CTT ATA TGT CGC GTC CAC GGA TCT CCC AAT ACG GTT ATT GCC TGG GAT TAC
CAT TGC GAA TAT ACA GCG CAG GTG CCT AGA GGG TTA TGC CAA TAA CGG ACC CTA ATG
296 V T L I C R V H G S P N T V I A W D Y
4398 ACC AAT CAA GTA TAC GAA TCC CGC TCC AAG CCG GTG AAA AGT CTG CAA AAG CAG CGC
TGG TTA GTT CAT ATG CTT AGG GCG AGG TTC GGC CAC TTT TCA GAC GTT TTC GTC GCG
315 T N Q V Y E S R S K P V K S L Q K Q R
4455 ATC TAC ATA GAA CTG TTG CGT GAG GAT GAA TCA AAG ATT CGA AAA TTC GGA CAT GAT
TAG ATG TAT CTT GAC AAC GCA CTC CTA CTT AGT TTC TAA GCT TTT AAG CCT GTA CTA
334 I Y I E L L R E D E S K I R K F G H D
4512 GTT TTT GTG TCA CGT CTA ACG ATA GTT AAT GCC CGT AAG AGC GAT GAG GGC GTC TAT
CAA AAA CAC AGT GCA GAT TGC TAT CAA TTA CGG GCA TTC TCG CTA CTC CCG CAG ATA
353 V F V S R L T I V N A R K S D E G V Y
4569 ACC TGT CTG GCG GAA AAT CCC GGT GGC AAG GAT TCG GTT CAC ATA AGT GTG GTG GTG
TGG ACA GAC CGC CTT TTA GGG CCA CCG TTC CTA AGC CAA GTG TAT TCA CAC CAC CAC
372 T C L A E N P G G K D S V H I S V V V
4626 CAA AAG GAT ATG GAA AGG ATT TCC CTC ATC GAC AGC AAC TTC TTT GCA ATA GTC TGC
GTT TTC CTA TAC CTT TCC TAA AGG GAG TAG CTG TCG TTG AAG AAA CGT TAT CAG ACG
391 Q K D M E R I S L I D S N F F A I V C
4683 CTT ATA GCC ATG GGT TTT CTC AGC ATG TCG ATT CTA TTT TCG TTG GTA ACA TGC CTA
GAA TAT CGG TAC CCA AAA GAG TCG TAC AGC TAA GAT AAA AGC AAC CAT TGT ACG GAT
410 L I A M G F L S M S I L F S L V T C L
4740 ATA TTT AAG AGA TTC AAG CAG TTC CAT CCC GGC CAG CAC ACT TAT TTG CAA CCC ACT
TAT AAA TTC TCT AAG TTC GTC AAG GTA GGG CCG GTC GTG TGA ATA AAC GTT GGG TGA
429 I F K R F K Q F H P G Q H T Y L Q P T
4797 AGT TTG CCC GTT CAG TCA CCT GGC AGT GAA GAA GCC ACC GCC ATC AGC GCC CTA AGT
TCA AAC GGG CAA GTC AGT GGA CCG TCA CTT CTT CGG TGG CGG TAG TCG CGG GAT TCA
448 S L P V Q S P G S E E A T A I S A L S
BamHI
4854 TCT GGA GTT ATA AGG GAA AGT AAA ATA GTG CTG GAT CCA TTA AGT GCT ATC AAT GAA
AGA CCT CAA TAT TCC CTT TCA TTT TAT CAC GAC CTA GGT AAT TCA CGA TAG TTA CTT
467 S G V I R E S K I V L D P L S A I N E
4911 CCA TCA AAT AAA AAT TAC ACT TTA TTT AAA ACA TCC AAT TCG AAT GGC AGC GAG TAT
GGT AGT TTA TTT TTA ATG TGA AAT AAA TTT TGT AGG TTA AGC TTA CCG TCG CTC ATA
486 P S N K N Y T L F K T S N S N G S E Y
4968 ATG CAC ACG AGA AAT TAT AAA GAC GTG AGA TTA AAC AGC AAC ACA TAT ACT GAA AAT
TAC GTG TGC TCT TTA ATA TTT CTG CAC TCT AAT TTG TCG TTG TGT ATA TGA CTT TTA
505 M H T R N Y K D V R L N S N T Y T E N
5025 CTA GAC AAT CAG GCG GAA TCC ATT TCA TCG CGG AAT CGG GAG CTC TAT TCA AAT ATA
GAT CTG TTA GTC CGC CTT AGG TAA AGT AGC GCC TTA GCC CTC GAG ATA AGT TTA TAT
524 L D N Q A E S I S S R N R E L Y S N I


attB2
V5 epitope
6xHis
SV40 Poly A


6156 TTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAA
AAAATGAACGAAATTTTTTGGAGGGTGTGGAGGGGGACTTGGACTTTGTATTTTACTTACGTTAACAACAACAATT
6232 CTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCA
GAACAAATAACGTCGAATATTACCAATGTTTATTTCGTTATCGTAGTGTTTAAAGTGTTTATTTCGTAAAAAAAGT
BamHI white gene
6308 CTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGGATCGGATCCACTAGAAGGCCTT
GACGTAAGATCAACACCAAACAGGTTTGAGTAGTTACATAGAATAGTACAGACCTAGCCTAGGTGATCTTCCGGAA
6384 AGTATGTATGTAAGTTAATAAAACCCTTTTTTGGAGAATGTAGATTTAAAAAAACATATTTTTTTTTTATTTTTTA
TCATACATACATTCAATTATTTTGGGAAAAAACCTCTTACATCTAAATTTTTTTGTATAAAAAAAAAATAAAAAAT
6460 CTGCACTGGACATCATTGAACTTATCTGATCAGTTTTAAATTTACTTCGATCCAAGGGTATTTGAAGTACCAGGTT
GACGTGACCTGTAGTAACTTGAATAGACTAGTCAAAATTTAAATGAAGCTAGGTTCCCATAAACTTCATGGTCCAA
6536 CTTTCGATTACCTCTCACTCAAAATGACATTCCACTCAAAGTCAGCGCTGTTTGCCTCCTTCTCTGTCCACAGAAA
GAAAGCTAATGGAGAGTGAGTTTTACTGTAAGGTGAGTTTCAGTCGCGACAAACGGAGGAAGAGACAGGTGTCTTT
6612 TATCGCCGTCTCTTTCGCCGCTGCGTCCGCTATCTCTTTCGCCACCGTTTGTAGCGTTACCTAGCGTCAATGTCCG
ATAGCGGCAGAGAAAGCGGCGACGCAGGCGATAGAGAAAGCGGTGGCAAACATCGCAATGGATCGCAGTTACAGGC
6688 CCTTCAGTTGCACTTTGTCAGCGGTTTCGTGACGAAGCTCCAAGCGGTTTACGCCATCAATTAAACACAAAGTGCT
GGAAGTCAACGTGAAACAGTCGCCAAAGCACTGCTTCGAGGTTCGCCAAATGCGGTAGTTAATTTGTGTTTCACGA
6764 GTGCCAAAACTCCTCTCGCTTCTTATTTTTGTTTGTTTTTTGAGTGATTGGGGTGGTGATTGGTTTTGGGTGGGTA
CACGGTTTTGAGGAGAGCGAAGAATAAAAACAAACAAAAAACTCACTAACCCCACCACTAACCAAAACCCACCCAT
6840 AGCAGGGGAAAGTGTGAAAAATCCCGGCAATGGGCCAAGAGGATCAGGAGCTATTAATTCGCGGAGGCAGCAAACA
TCGTCCCCTTTCACACTTTTTAGGGCCGTTACCCGGTTCTCCTAGTCCTCGATAATTAAGCGCCTCCGTCGTTTGT
6916 CCCATCTGCCGAGCATCTGAACAATGTGAGTAGTACATGTGCATACATCTTAAGTTCACTTGATCTATAGGAACTG
GGGTAGACGGCTCGTAGACTTGTTACACTCATCATGTACACGTATGTAGAATTCAAGTGAACTAGATATCCTTGAC
6992 CGATTGCAACATCAAATTGTCTGCGGCGTGAGAACTGCGACCCACAAAAATCCCAAACCGCAATCGCACAAACAAA
GCTAACGTTGTAGTTTAACAGACGCCGCACTCTTGACGCTGGGTGTTTTTAGGGTTTGGCGTTAGCGTGTTTGTTT
7068 TAGTGACACGAAACAGATTATTCTGGTAGCTGTGCTCGCTATATAAGACAATTTTTAAGATCATATCATGATCAAG
ATCACTGTGCTTTGTCTAATAAGACCATCGACACGAGCGATATATTCTGTTAAAAATTCTAGTATAGTACTAGTTC
7144 ACATCTAAAGGCATTCATTTTCGACTACATTCTTTTTTACAAAAAATATAACAACCAGATATTTTAAGCTGATCCT
TGTAGATTTCCGTAAGTAAAAGCTGATGTAAGAAAAAATGTTTTTTATATTGTTGGTCTATAAAATTCGACTAGGA
7220 AGATGCACAAAAAATAAATAAAAGTATAAACCTACTTCGTAGGATACTTCGTTTTGTTCGGGGTTAGATGAGCATA
TCTACGTGTTTTTTATTTATTTTCATATTTGGATGAAGCATCCTATGAAGCAAAACAAGCCCCAATCTACTCGTAT
7296 ACGCTTGTAGTTGATATTTGAGATCCCCTATCATTGCAGGGTGACAGCGGAGCGGCTTCGCAGAGCTGCATTAACC
TGCGAACATCAACTATAAACTCTAGGGGATAGTAACGTCCCACTGTCGCCTCGCCGAAGCGTCTCGACGTAATTGG
7372 AGGGCTTCGGGCAGGCCAAAAACTACGGCACGCTCCTGCCACCCAGTCCGCCGGAGGACTCCGGTTCAGGGAGCGG
TCCCGAAGCCCGTCCGGTTTTTGATGCCGTGCGAGGACGGTGGGTCAGGCGGCCTCCTGAGGCCAAGTCCCTCGCC

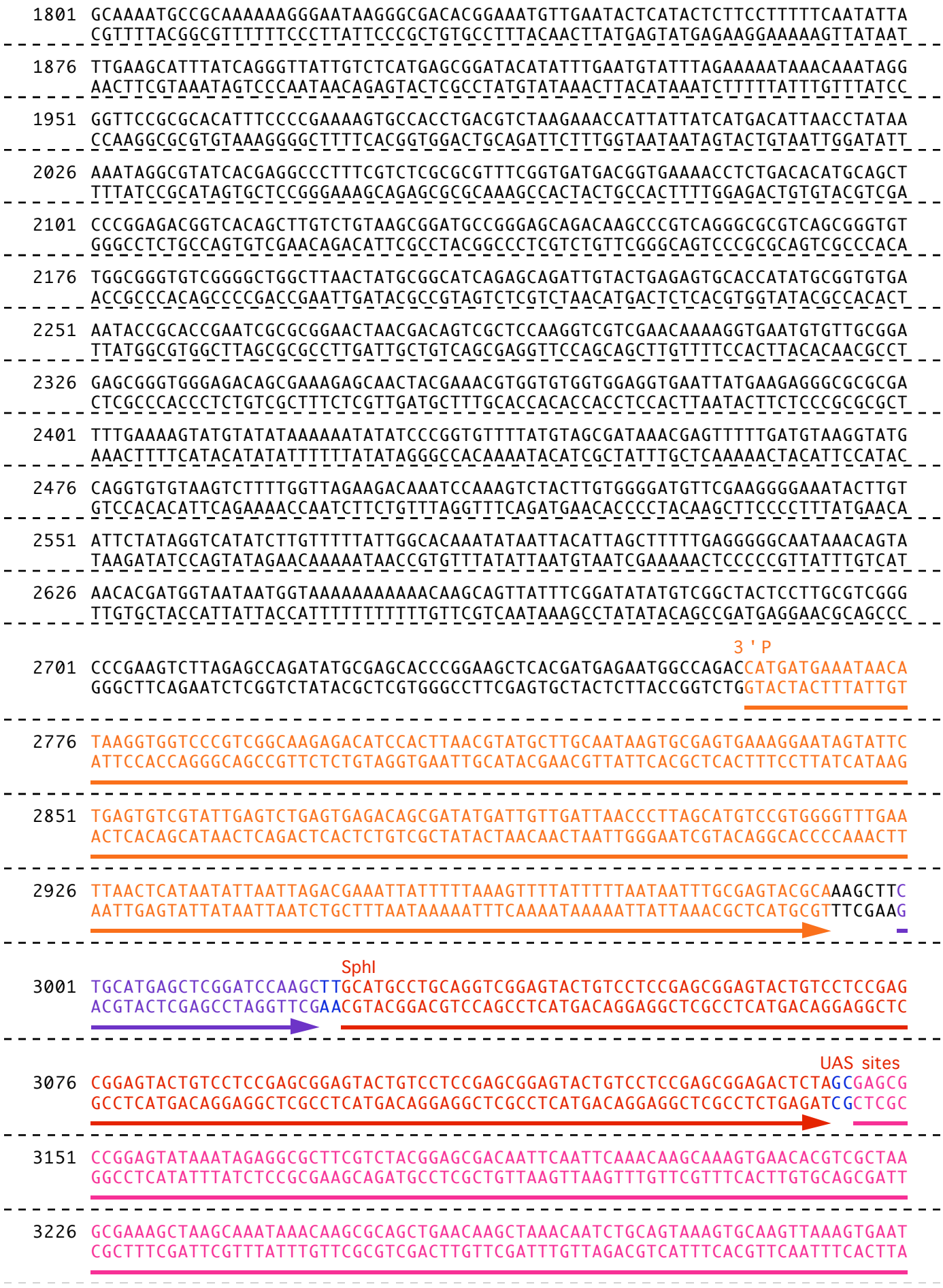

7448 CCAACTAGCCGAGAACCTCACCTATGCCTGGCACAATATGGACATCTTTGGGGCGGTCAATCAGCCGGGCTCCGGA
GGTTGATCGGCTCTTGGAGTGGATACGGACCGTGTTATACCTGTAGAAACCCCGCCAGTTAGTCGGCCCGAGGCCT
7524 TGGCGGCAGCTGGTCAACCGGACACGCGGACTATTCTGCAACGAGCGACACATACCGGCGCCCAGGAAACATTTGC
ACCGCCGTCGACCAGTTGGCCTGTGCGCCTGATAAGACGTTGCTCGCTGTGTATGGCCGCGGGTCCTTTGTAAACG
7600 TCAAGAACGGTGAGTTTCTATTCGCAGTCGGCTGATCTGTGTGAAATCTTAATAAAGGGTCCAATTACCAATTTGA
AGTTCTTGCCACTCAAAGATAAGCGTCAGCCGACTAGACACACTTTAGAATTATTTCCCAGGTTAATGGTTAAACT
7676 AACTCAGTTTGCGGCGTGGCCTATCCGGGCGAACTTTTGGCCGTGATGGGCAGTTCCGGTGCCGGAAAGACGACCC
TTGAGTCAAACGCCGCACCGGATAGGCCCGCTTGAAAACCGGCACTACCCGTCAAGGCCACGGCCTTTCTGCTGGG
7752 TGCTGAATGCCCTTGCCTTTCGATCGCCGCAGGGCATCCAAGTATCGCCATCCGGGATGCGACTGCTCAATGGCCA
ACGACTTACGGGAACGGAAAGCTAGCGGCGTCCCGTAGGTTCATAGCGGTAGGCCCTACGCTGACGAGTTACCGGT
7828 ACCTGTGGACGCCAAGGAGATGCAGGCCAGGTGCGCCTATGTCCAGCAGGATGACCTCTTTATCGGCTCCCTAACG
TGGACACCTGCGGTTCCTCTACGTCCGGTCCACGCGGATACAGGTCGTCCTACTGGAGAAATAGCCGAGGGATTGC
7904 GCCAGGGAACACCTGATTTTCCAGGCCATGGTGCGGATGCCACGACATCTGACCTATCGGCAGCGAGTGGCCCGCG
CGGTCCCTTGTGGACTAAAAGGTCCGGTACCACGCCTACGGTGCTGTAGACTGGATAGCCGTCGCTCACCGGGCGC
7980 TGGATCAGGTGATCCAGGAGCTTTCGCTCAGCAAATGTCAGCACACGATCATCGGTGTGCCCGGCAGGGTGAAAGG
ACCTAGTCCACTAGGTCCTCGAAAGCGAGTCGTTTACAGTCGTGTGCTAGTAGCCACACGGGCCGTCCCACTTTCC
8056 TCTGTCCGGCGGAGAAAGGAAGCGTCTGGCATTCGCCTCCGAGGCACTAACCGATCCGCCGCTTCTGATCTGCGAT
AGACAGGCCGCCTCTTTCCTTCGCAGACCGTAAGCGGAGGCTCCGTGATTGGCTAGGCGGCGAAGACTAGACGCTA
8132 GAGCCCACCTCCGGACTGGACTCATTTACCGCCCACAGCGTCGTCCAGGTGCTGAAGAAGCTGTCGCAGAAGGGCA
CTCGGGTGGAGGCCTGACCTGAGTAAATGGCGGGTGTCGCAGCAGGTCCACGACTTCTTCGACAGCGTCTTCCCGT
8208 AGACCGTCATCCTGACCATTCATCAGCCGTCTTCCGAGCTGTTTGAGCTCTTTGACAAGATCCTTCTGATGGCCGA
TCTGGCAGTAGGACTGGTAAGTAGTCGGCAGAAGGCTCGACAAACTCGAGAAACTGTTCTAGGAAGACTACCGGCT
8284 GGGCAGGGTAGCTTTCTTGGGCACTCCCAGCGAAGCCGTCGACTTCTTTTCCTAGTGAGTTCGATGTGTTTATTAA
CCCGTCCCATCGAAAGAACCCGTGAGGGTCGCTTCGGCAGCTGAAGAAAAGGATCACTCAAGCTACACAAATAATT
8360 GGGTATCTAGCATTACATTACATCTCAACTCCTATCCAGCGTGGGTGCCCAGTGTCCTACCAACTACAATCCGGCG
CCCATAGATCGTAATGTAATGTAGAGTTGAGGATAGGTCGCACCCACGGGTCACAGGATGGTTGATGTTAGGCCGC
8436 GACTTTTACGTACAGGTGTTGGCCGTTGTGCCCGGACGGGAGATCGAGTCCCGTGATCGGATCGCCAAGATATGCG
CTGAAAATGCATGTCCACAACCGGCAACACGGGCCTGCCCTCTAGCTCAGGGCACTAGCCTAGCGGTTCTATACGC
8512 ACAATTTTGCTATTAGCAAAGTAGCCCGGGATATGGAGCAGTTGTTGGCCACCAAAAATTTGGAGAAGCCACTGGA
TGTTAAAACGATAATCGTTTCATCGGGCCCTATACCTCGTCAACAACCGGTGGTTTTTAAACCTCTTCGGTGACCT
8588 GCAGCCGGAGAATGGGTACACCTACAAGGCCACCTGGTTCATGCAGTTCCGGGCGGTCCTGTGGCGATCCTGGCTG
CGTCGGCCTCTTACCCATGTGGATGTTCCGGTGGACCAAGTACGTCAAGGCCCGCCAGGACACCGCTAGGACCGAC
8664 TCGGTGCTCAAGGAACCACTCCTCGTAAAAGTGCGACTTATTCAGACAACGGTGAGTGGTTCCAGTGGAAACAAAT
AGCCACGAGTTCCTTGGTGAGGAGCATTTTCACGCTGAATAAGTCTGTTGCCACTCACCAAGGTCACCTTTGTTTA
8740 GATATAACGCTTACAATTCTTGGAAACAAATTCGCTAGATTTTAGTTAGAATTGCCTGATTCCACACCCTTCTTAG
CTATATTGCGAATGTTAAGAACCTTTGTTTAAGCGATCTAAAATCAATCTTAACGGACTAAGGTGTGGGAAGAATC

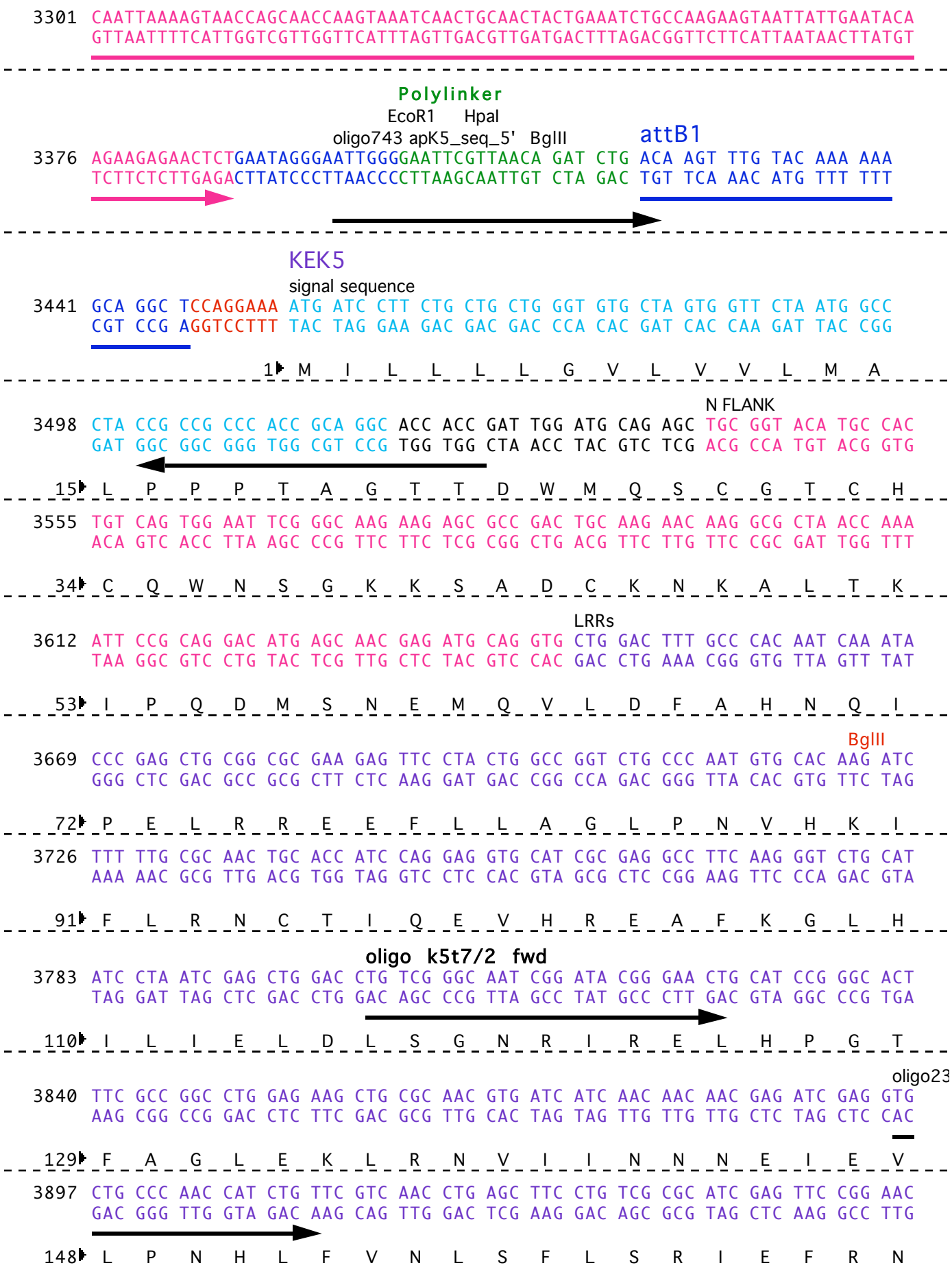

8816 TTTTTTTCAATGAGATGTATAGTTTATAGTTTTGCAGAAAATAAATAAATTTCATTTAACTCGCGAACATGTTGAA
AAAAAAAGTTACTCTACATATCAAATATCAAAACGTCTTTTATTTATTTAAAGTAAATTGAGCGCTTGTACAACTT
8892 GATATGAATATTAATGAGATGCGAGTAACATTTTAATTTGCAGATGGTTGCCATCTTGATTGGCCTCATCTTTTTG
CTATACTTATAATTACTCTACGCTCATTGTAAAATTAAACGTCTACCAACGGTAGAACTAACCGGAGTAGAAAAAC
8968 GGCCAACAACTCACGCAAGTGGGCGTGATGAATATCAACGGAGCCATCTTCCTCTTCCTGACCAACATGACCTTTC
CCGGTTGTTGAGTGCGTTCACCCGCACTACTTATAGTTGCCTCGGTAGAAGGAGAAGGACTGGTTGTACTGGAAAG
9044 AAAACGTCTTTGCCACGATAAATGTAAGTCTTGTTTAGAATACATTTGCATATTAATAATTTACTAACTTTCTAAT
TTTTGCAGAAACGGTGCTATTTACATTCAGAACAAATCTTATGTAAACGTATAATTATTAAATGATTGAAAGATTA
9120 GAATCGATTCGATTTAGGTGTTCACCTCAGAGCTGCCAGTTTTTATGAGGGAGGCCCGAAGTCGACTTTATCGCTG
CTTAGCTAAGCTAAATCCACAAGTGGAGTCTCGACGGTCAAAAATACTCCCTCCGGGCTTCAGCTGAAATAGCGAC
9196 TGACACATACTTTCTGGGCAAAACGATTGCCGAATTACCGCTTTTTCTCACAGTGCCACTGGTCTTCACGGCGATT
ACTGTGTATGAAAGACCCGTTTTGCTAACGGCTTAATGGCGAAAAAGAGTGTCACGGTGACCAGAAGTGCCGCTAA
9272 GCCTATCCGATGATCGGACTGCGGGCCGGAGTGCTGCACTTCTTCAACTGCCTGGCGCTGGTCACTCTGGTGGCCA
CGGATAGGCTACTAGCCTGACGCCCGGCCTCACGACGTGAAGAAGTTGACGGACCGCGACCAGTGAGACCACCGGT
9348 ATGTGTCAACGTCCTTCGGATATCTAATATCCTGCGCCAGCTCCTCGACCTCGATGGCGCTGTCTGTGGGTCCGCC
TACACAGTTGCAGGAAGCCTATAGATTATAGGACGCGGTCGAGGAGCTGGAGCTACCGCGACAGACACCCAGGCGG
9424 GGTTATCATACCATTCCTGCTCTTTGGCGGCTTCTTCTTGAACTCGGGCTCGGTGCCAGTATACCTCAAATGGTTG
CCAATAGTATGGTAAGGACGAGAAACCGCCGAAGAAGAACTTGAGCCCGAGCCACGGTCATATGGAGTTTACCAAC
9500 TCGTACCTCTCATGGTTCCGTTACGCCAACGAGGGTCTGCTGATTAACCAATGGGCGGACGTGGAGCCGGGCGAAA
AGCATGGAGAGTACCAAGGCAATGCGGTTGCTCCCAGACGACTAATTGGTTACCCGCCTGCACCTCGGCCCGCTTT
9576 TTAGCTGCACATCGTCGAACACCACGTGCCCCAGTTCGGGCAAGGTCATCCTGGAGACGCTTAACTTCTCCGCCGC
AATCGACGTGTAGCAGCTTGTGGTGCACGGGGTCAAGCCCGTTCCAGTAGGACCTCTGCGAATTGAAGAGGCGGCG
9652 CGATCTGCCGCTGGACTACGTGGGTCTGGCCATTCTCATCGTGAGCTTCCGGGTGCTCGCATATCTGGCTCTAAGA
GCTAGACGGCGACCTGATGCACCCAGACCGGTAAGAGTAGCACTCGAAGGCCCACGAGCGTATAGACCGAGATTCT
9728 CTTCGGGCCCGACGCAAGGAGTAGCCGACATATATCCGAAATAACTGCTTGTTTTTTTTTTTACCATTATTACCAT
GAAGCCCGGGCTGCGTTCCTCATCGGCTGTATATAGGCTTTATTGACGAACAAAAAAAAAAAATGGTAATAATGGTA
9804 CGTGTTTACTGTTTATTGCCCCCTCAAAAAGCTAATGTAATTATATTTGTGCCAATAAAAACAAGATATGACCTAT
GCACAAATGACAAATAACGGGGGAGTTTTTCGATTACATTAATATAAACACGGTTATTTTTGTTCTATACTGGATA
9880 AGAATACAAGTATTTCCCCTTCGAACATCCCCACAAGTAGACTTTGGATTTGTCTTCTAACCAAAAGACTTACACA
TCTTATGTTCATAAAGGGGAAGCTTGTAGGGGTGTTCATCTGAAACCTAAACAGAAGATTGGTTTTCTGAATGTGT
9956 CCTGCATACCTTACATCAAAAACTCGTTTATCGCTACATAAAACACCGGGATATATTTTTTATATACATACTTTTC
GGACGTATGGAATGTAGTTTTTGAGCAAATAGCGATGTATTTTGTGGCCCTATATAAAAAATATATGTATGAAAAG
10032 AAATCGCGCGCCCTCTTCATAATTCACCTCCACCACACCACGTTTCGTAGTTGCTCTTTCGCTGTCTCCCACCCGC
TTTAGCGCGCGGGAGAAGTATTAAGTGGAGGTGGTGTGGTGCAAAGCATCAACGAGAAAGCGACAGAGGGTGGGCG
10108 TCTCCGCAACACATTCACCTTTTGTTCGACGACCTTGGAGCGACTGTCGTTAGTTCCGCGCGATTCGGTTCGCTCA
AGAGGCGTTGTGTAAGTGGAAAACAAGCTGCTGGAACCTCGCTGACAGCAATCAAGGCGCGCTAAGCCAAGCGAGT


10184 AATGGTTCCGAGTGGTTCATTTCGTCTCAATAGAAATTAGTAATAAATATTTGTATGTACAATTTATTTGCTCCAA
TTACCAAGGCTCACCAAGTAAAGCAGAGTTATCTTTAATCATTATTTATAAACATACATGTTAAATAAACGAGGTT
10260 TATATTTGTATATATTTCCCTCACAGCTATATTTATTCTAATTTAATATTATGACTTTTTAAGGTAATTTTTTGTG
ATATAAACATATATAAAGGGAGTGTCGATATAAATAAGATTAAATTATAATACTGAAAAATTCCATTAAAAAACAC
10336 ACCTGTTCGGAGTGATTAGCGTTACAATTTGAACTGAAAGTGACATCCAGTGTTTGTTCCTTGTGTAGATGCATCT
TGGACAAGCCTCACTAATCGCAATGTTAAACTTGACTTTCACTGTAGGTCACAAACAAGGAACACATCTACGTAGA
10412 CAAAAAAATGGTGGGCATAATAGTGTTGTTTATATATATCAAAAATAACAACTATAATAATAAGAATACATTTAAT
GTTTTTTTACCACCCGTATTATCACAACAAATATATATAGTTTTTATTGTTGATATTATTATTCTTATGTAAATTA
10488 TTAGAAAATGCTTGGATTTCACTGGAACTAGAATTAATTCGGCTGCTGCTCTAAACGACGCATTTCGTACTCCAAA
AATCTTTTACGAACCTAAAGTGACCTTGATCTTAATTAAGCCGACGACGAGATTTGCTGCGTAAAGCATGAGGTTT
10564 GTACGAATTTTTTCCCTCAAGCTCTTATTTTCATTAAACAATGAACAGGACCTAACGCACAGTCACGTTATTGTTT
CATGCTTAAAAAAGGGAGTTCGAGAATAAAAGTAATTTGTTACTTGTCCTGGATTGCGTGTCAGTGCAATAACAAA
10640 ACATAAATGATTTTTTTTACTATTCAAACTTACTCTGTTTGTGTACTCCCACTGGTATAGCCTTCTTTTATCTTTT
TGTATTTACTAAAAAAAATGATAAGTTTGAATGAGACAAACACATGAGGGTGACCATATCGGAAGAAAATAGAAAA
10716 CTGGTTCAGGCTCTATCACTTTACTAGGTACGGCATCTGCGTTGAGTCGCCTCCTTTTAAATGTCTGACCTTTTGC
GACCAAGTCCGAGATAGTGAAATGATCCATGCCGTAGACGCAACTCAGCGGAGGAAAATTTACAGACTGGAAAACG
10792 AGGTGCAGCCTTCCACTGCGAATCATTAAAGTGGGTATCACAAATTTGGGAGTTTTCACCAAGGCTGCACCCAAGG
TCCACGTCGGAAGGTGACGCTTAGTAATTTCACCCATAGTGTTTAAACCCTCAAAAGTGGTTCCGACGTGGGTTCC
10868 CTCTGCTCCCACAATTTTCTCTTAATAGCACACTTCGGCACGTGAATTAATTTTACTCCAGTCACAGCTTTGCAGC
GAGACGAGGGTGTTAAAAGAGAATTATCGTGTGAAGCCGTGCACTTAATTAAAATGAGGTCAGTGTCGAAACGTCG
10944 AAAATTTGCAATATTTCATTTTTTTTTATTCCACGTAAGGGTTAATGTTTTCAAAAAAAAATTCGTCCGCACACAA
TTTTAAACGTTATAAAGTAAAAAAAAATAAGGTGCATTCCCAATTACAAAAGTTTTTTTTTAAGCAGGCGTGTGTT
11020 CCTTTCCTCTCAACAAGCAAACGTGCACTGAATTTAAGTGTATACTTCGGTAAGCTTCGGCTATCGACGGGACCAC
GGAAAGGAGAGTTGTTCGTTTGCACGTGACTTAAATTCACATATGAAGCCATTCGAAGCCGATAGCTGCCCTGGTG
11096 CTTATGTTATTTCATCATG
GAATACAATAAAGTAGTAC

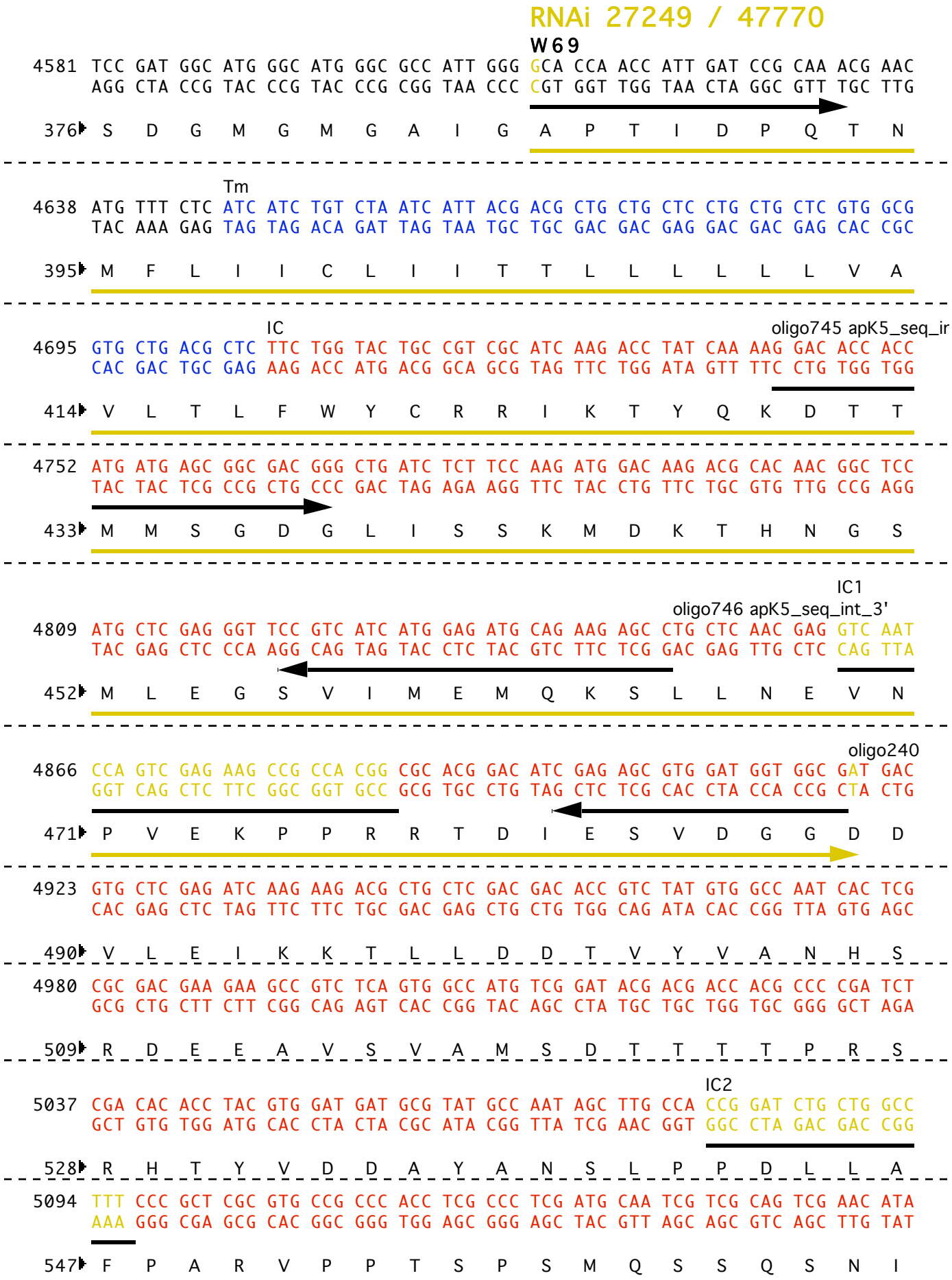
UAS Kek5 GFP

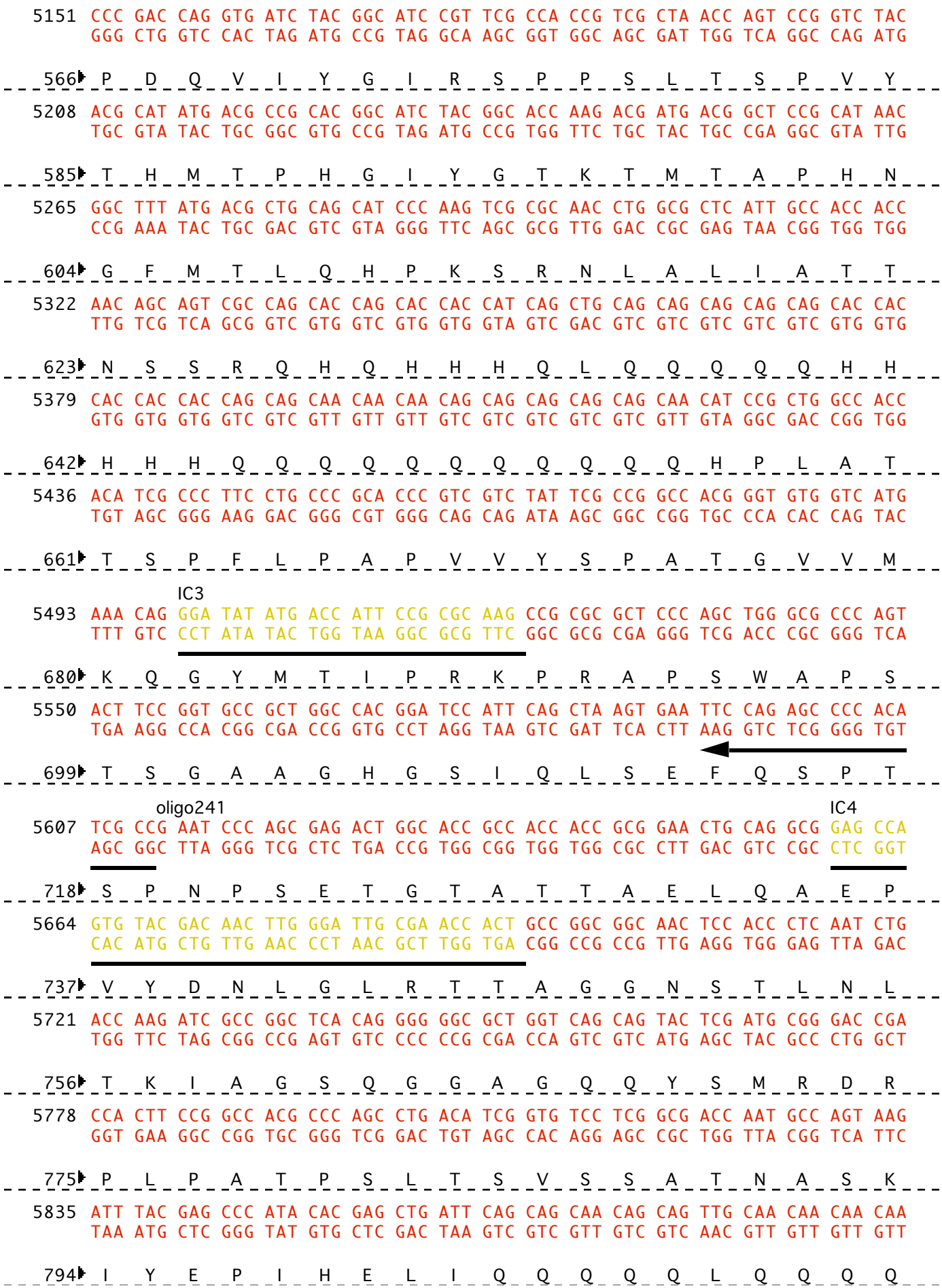
```
     1 GGCCAGACCCACGTAGTCCAGCGGCAGATCGGCGGCGGAGAAGTTAAGCGTCTCCAGGATGACCTTGCCCGAACT
       CCGGTCTGGGTGCATCAGGTCGCCGTCTAGCCGCCGCCTCTTCAATTCGCAGAGGTCCTACTGGAACGGGCTTGA

    76 GGGGCACGTGGTGTTCGACGATGTGCAGCTAATTTCGCCCGGCTCCACGTCCGCCCATTGGTTAATCAGCAGACC
       CCCCGTGCACCACAAGCTGCTACACGTCGATTAAAGCGGGCCGAGGTGCAGGCGGGTAACCAATTAGTCGTCTGG

   151 CTCGTTGGCGTAACGGAACCATGAGAGGTACGACAACCATTTGAGGTATACTGGCACCGAGCCCGAGTTCAAGAA
       GAGCAACCGCATTGCCTTGGTACTCTCCATGCTGTTGGTAAACTCCATATGACCGTGGCTCGGGCTCAAGTTCTT

   226 GAAGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA
       CTTCCGCAAAAAGGTATCCGAGGCGGGGGGACTGCTCGTAGTGTTTTTAGCTGCGAGTTCAGTCTCCACCGCTTT

   301 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC
       GGGCTGTCCTGATATTTCTATGGTCCGCAAAGGGGGACCTTCGAGGGAGCACGCGAGAGGACAAGGCTGGGACGG

   376 GCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCT
       CGAATGGCCTATGGACAGGCGGAAAGAGGGAAGCCCTTCGCACCGCGAAAGAGTTACGAGTGCGACATCCATAGA

   451 CAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTT
       GTCAAGCCACATCCAGCAAGCGAGGTTCGACCCGACACACGTGCTTGGGGGGCAAGTCGGGCTGGCGACGCGGAA

   526 ATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG
       TAGGCCATTGATAGCAGAACTCAGGTTGGGCCATTCTGTGCTGAATAGCGGTGACCGTCGTCGGTGACCATTGTC

   601 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAG
       CTAATCGTCTCGCTCCATACATCCGCCACGATGTCTCAAGAACTTCACCACCGGATTGATGCCGATGTGATCTTC

   676 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA
       CTGTCATAAACCATAGACGCGAGACGACTTCGGTCAATGGAAGCCTTTTTCTCAACCATCGAGAACTAGGCCGTT

   751 ACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGA
       TGTTTGGTGGCGACCATCGCCACCAAAAAAACAAACGTTCGTCGTCTAATGCGCGTCTTTTTTTCCTAGAGTTCT

   826 AGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG
       TCTAGGAAACTAGAAAAGATGCCCCAGACTGCGAGTCACCTTGCTTTTGAGTGCAATTCCCTAAAACCAGTACTC

   901 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGA
       TAATAGTTTTTCCTAGAAGTGGATCTAGGAAAATTTAATTTTTACTTCAAAATTTAGTTAGATTTCATATATACT

   976 GTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATC
       CATTTGAACCAGACTGTCAATGGTTACGAATTAGTCACTCCGTGGATAGAGTCGCTAGACAGATAAAGCAAGTAG

  1051 CATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAAT
       GTATCAACGGACTGAGGGGCAGCACATCTATTGATGCTATGCCCTCCCGAATGGTAGACCGGGGTCACGACGTTA

  1126 GATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG
       CTATGGCGCTCTGGGTGCGAGTGGCCGAGGTCTAAATAGTCGTTATTTGGTCGGTCGGCCTTCCCGGCTCGCGTC

  1201 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCC
       TTCACCAGGACGTTGAAATAGGCGGAGGTAGGTCAGATAATTAACAACGGCCCTTCGATCTCATTCATCAAGCGG

  1276 AGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTC
       TCAATTATCAAACGCGTTGCAACAACGGTAACGATGTCCGTAGCACCACAGTGCGAGCAGCAAACCATACCGAAG

  1351 ATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTT
       TAAGTCGAGGCCAAGGGTTGCTAGTTCCGCTCAATGTACTAGGGGGTACAACACGTTTTTTCGCCAATCGAGGAA

  1426 CGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC
       GCCAGGAGGCTAGCAACAGTCTTCATTCAACCGGCGTCACAATAGTGAGTACCAATACCGTCGTGACGTATTAAG

  1501 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG
       AGAATGACAGTACGGTAGGCATTCTACGAAAAGACACTGACCACTCATGAGTTGGTTCAGTAAGACTCTTATCAC

  1576 TATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGT
       ATACGCCGCTGGCTCAACGAGAACGGGCCGCAGTTGTGCCCTATTATGGCGCGGTGTATCGTCTTGAAATTTTCA

  1651 GCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTA
       CGAGTAGTAACCTTTTGCAAGAAGCCCCGCTTTTGAGAGTTCCTAGAATGGCGACAACTCTAGGTCAAGCTACAT

  1726 ACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG
       TGGGTGAGCACGTGGGTTGACTAGAAGTCGTAGAAAATGAAAGTGGTCGCAAAGACCCACTCGTTTTTGTCCTTC
```

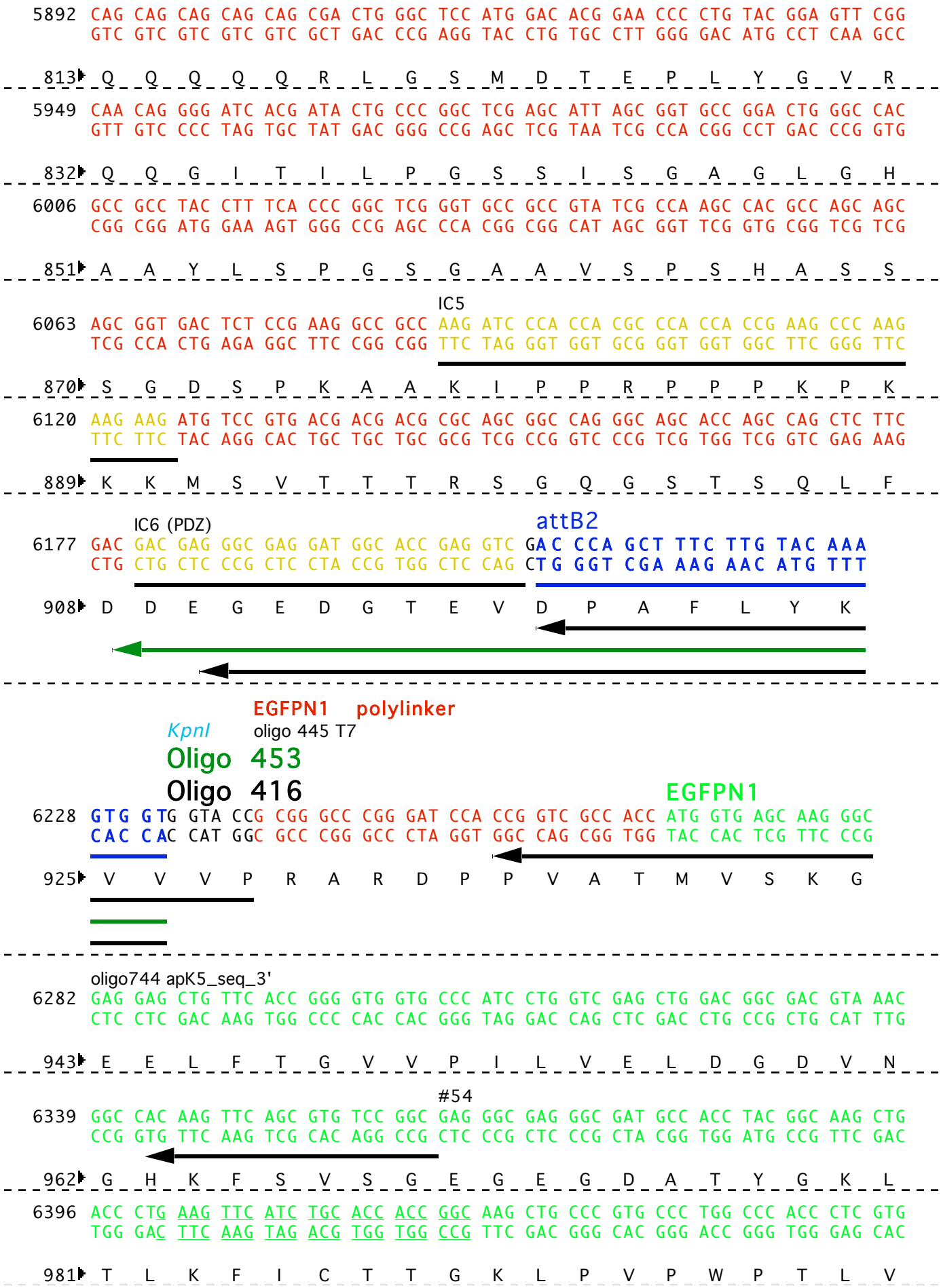

1801 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTA
CGTTTTACGGCGTTTTTTCCCTTATTCCCGCTGTGCCTTTACAACTTATGAGTATGAGAAGGAAAAAGTTATAAT
1876 TTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGG
AACTTCGTAAATAGTCCCAATAACAGAGTACTCGCCTATGTATAAACTTACATAAATCTTTTTATTTGTTTATCC
1951 GGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAA
CCAAGGCGCGTGTAAAGGGGCTTTTCACGGTGGACTGCAGATTCTTTGGTAATAATAGTACTGTAATTGGATATT
2026 AAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCT
TTTATCCGCATAGTGCTCCGGGAAAGCAGAGCGCGCAAAGCCACTACTGCCACTTTTGGAGACTGTGTACGTCGA
2101 CCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGT
GGGCCTCTGCCAGTGTCGAACAGACATTCGCCTACGGCCCTCGTCTGTTCGGGCAGTCCCGCGCAGTCGCCCACA
2176 TGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGA
ACCGCCCACAGCCCCGACCGAATTGATACGCCGTAGTCTCGTCTAACATGACTCTCACGTGGTATACGCCACACT
2251 AATACCGCACCGAATCGCGCGGAACTAACGACAGTCGCTCCAAGGTCGTCGAACAAAAGGTGAATGTGTTGCGGA
TTATGGCGTGGCTTAGCGCGCCTTGATTGCTGTCAGCGAGGTTCCAGCAGCTTGTTTTCCACTTACACAACGCCT
2326 GAGCGGGTGGGAGACAGCGAAAGAGCAACTACGAAACGTGGTGTGGTGGAGGTGAATTATGAAGAGGGCGCGCGA
CTCGCCCACCCTCTGTCGCTTTCTCGTTGATGCTTTGCACCACACCACCTCCACTTAATACTTCTCCCGCGCGCT
2401 TTTGAAAAGTATGTATATAAAAAATATATCCCGGTGTTTTATGTAGCGATAAACGAGTTTTTGATGTAAGGTATG
AAACTTTTCATACATATATTTTTTATATAGGGCCACAAAATACATCGCTATTTGCTCAAAAACTACATTCCATAC
2476 CAGGTGTGTAAGTCTTTTGGTTAGAAGACAAATCCAAAGTCTACTTGTGGGGATGTTCGAAGGGGAAATACTTGT
GTCCACACATTCAGAAAACCAATCTTCTGTTTAGGTTTCAGATGAACACCCCTACAAGCTTCCCCTTTATGAACA
2551 ATTCTATAGGTCATATCTTGTTTTTATTGGCACAAATATAATTACATTAGCTTTTTGAGGGGGCAATAAACAGTA
TAAGATATCCAGTATAGAACAAAAATAACCGTGTTTATATTAATGTAATCGAAAAACTCCCCCGTTATTTGTCAT
2626 AACACGATGGTAATAATGGTAAAAAAAAAAACAAGCAGTTATTTCGGATATATGTCGGCTACTCCTTGCGTCGGG
TTGTGCTACCATTATTACCATTTTTTTTTTTGTTCGTCAATAAAGCCTATATACAGCCGATGAGGAACGCAGCCC
3 ' P
2701 CCCGAAGTCTTAGAGCCAGATATGCGAGCACCCGGAAGCTCACGATGAGAATGGCCAGACCATGATGAAATAACA
GGGCTTCAGAATCTCGGTCTATACGCTCGTGGGCCTTCGAGTGCTACTCTTACCGGTCTGGTACTACTTTATTGT
2776 TAAGGTGGTCCCGTCGGCAAGAGACATCCACTTAACGTATGCTTGCAATAAGTGCGAGTGAAAGGAATAGTATTC
ATTCCACCAGGGCAGCCGTTCTCTGTAGGTGAATTGCATACGAACGTTATTCACGCTCACTTTCCTTATCATAAG
2851 TGAGTGTCGTATTGAGTCTGAGTGAGACAGCGATATGATTGTTGATTAACCCTTAGCATGTCCGTGGGGTTTGAA
ACTCACAGCATAACTCAGACTCACTCTGTCGCTATACTAACAACTAATTGGGAATCGTACAGGCACCCCAAACTT
2926 TTAACTCATAATATTAATTAGACGAAATTATTTTTAAAGTTTTATTTTTAATAATTTGCGAGTACGCAAAGCTTC
AATTGAGTATTATAATTAATCTGCTTTAATAAAAATTTCAAAATAAAAATTATTAAACGCTCATGCGTTTCGAAG
SphI
3001 TGCATGAGCTCGGATCCAAGCTTGCATGCCTGCAGGTCGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAG
ACGTACTCGAGCCTAGGTTCGAACGTACGGACGTCCAGCCTCATGACAGGAGGCTCGCCTCATGACAGGAGGCTC
UAS sites
3076 CGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAGCGGAGACTCTAGCGAGCG
GCCTCATGACAGGAGGCTCGCCTCATGACAGGAGGCTCGCCTCATGACAGGAGGCTCGCCTCTGAGATCGCTCGC
3151 CCGGAGTATAAATAGAGGCGCTTCGTCTACGGAGCGACAATTCAATTCAAACAAGCAAAGTGAACACGTCGCTAA
GGCCTCATATTTATCTCCGCGAAGCAGATGCCTCGCTGTTAAGTTAAGTTTGTTCGTTTCACTTGTGCAGCGATT
3226 GCGAAAGCTAAGCAAATAAACAAGCGCAGCTGAACAAGCTAAACAATCTGCAGTAAAGTGCAAGTTAAAGTGAAT
CGCTTTCGATTCGTTTATTTGTTCGCGTCGACTTGTTCGATTTGTTAGACGTCATTTCACGTTCAATTTCACTTA


3301 CAATTAAAAGTAACCAGCAACCAAGTAAATCAACTGCAACTACTGAAATCTGCCAAGAAGTAATTATTGAATACA
GTTAATTTTCATTGGTCGTTGGTTCATTTAGTTGACGTTGATGACTTTAGACGGTTCTTCATTAATAACTTATGT
Polylinker
EcoR1 HpaI
oligo743 apK5_seq_5' BglII
attB1
3376 AGAAGAGAACTCTGAATAGGGAATTGGGGAATTCGTTAACA GAT CTG ACA AGT TTG TAC AAA AAA
TCTTCTCTTGAGACTTATCCCTTAACCCCTTAAGCAATTGT CTA GAC TGT TCA AAC ATG TTT TTT
KEK5
signal sequence
3441 GCA GGC TCCAGGAAA ATG ATC CTT CTG CTG CTG GGT GTG CTA GTG GTT CTA ATG GCC
CGT CCG AGGTCCTTT TAC TAG GAA GAC GAC GAC CCA CAC GAT CAC CAA GAT TAC CGG
1 M I L L L L G V L V V L M A
N FLANK
3498 CTA CCG CCG CCC ACC GCA GGC ACC ACC GAT TGG ATG CAG AGC TGC GGT ACA TGC CAC
GAT GGC GGC GGG TGG CGT CCG TGG TGG CTA ACC TAC GTC TCG ACG CCA TGT ACG GTG
15 L P P P T A G T T D W M Q S C G T C H
3555 TGT CAG TGG AAT TCG GGC AAG AAG AGC GCC GAC TGC AAG AAC AAG GCG CTA ACC AAA
ACA GTC ACC TTA AGC CCG TTC TTC TCG CGG CTG ACG TTC TTG TTC CGC GAT TGG TTT
34 C Q W N S G K K S A D C K N K A L T K
LRRs
3612 ATT CCG CAG GAC ATG AGC AAC GAG ATG CAG GTG CTG GAC TTT GCC CAC AAT CAA ATA
TAA GGC GTC CTG TAC TCG TTG CTC TAC GTC CAC GAC CTG AAA CGG GTG TTA GTT TAT
53 I P Q D M S N E M Q V L D F A H N Q I
BglII
3669 CCC GAG CTG CGG CGC GAA GAG TTC CTA CTG GCC GGT CTG CCC AAT GTG CAC AAG ATC
GGG CTC GAC GCC GCG CTT CTC AAG GAT GAC CGG CCA GAC GGG TTA CAC GTG TTC TAG
72 P E L R R E E F L L A G L P N V H K I
3726 TTT TTG CGC AAC TGC ACC ATC CAG GAG GTG CAT CGC GAG GCC TTC AAG GGT CTG CAT
AAA AAC GCG TTG ACG TGG TAG GTC CTC CAC GTA GCG CTC CGG AAG TTC CCA GAC GTA
91 F L R N C T I Q E V H R E A F K G L H
oligo k5t7/2 fwd
3783 ATC CTA ATC GAG CTG GAC CTG TCG GGC AAT CGG ATA CGG GAA CTG CAT CCG GGC ACT
TAG GAT TAG CTC GAC CTG GAC AGC CCG TTA GCC TAT GCC CTT GAC GTA GGC CCG TGA
110 I L I E L D L S G N R I R E L H P G T
oligo23
3840 TTC GCC GGC CTG GAG AAG CTG CGC AAC GTG ATC ATC AAC AAC AAC GAG ATC GAG GTG
AAG CGG CCG GAC CTC TTC GAC GCG TTG CAC TAG TAG TTG TTG TTG CTC TAG CTC CAC
129 F A G L E K L R N V I I N N N E I E V
3897 CTG CCC AAC CAT CTG TTC GTC AAC CTG AGC TTC CTG TCG CGC ATC GAG TTC CGG AAC
GAC GGG TTG GTA GAC AAG CAG TTG GAC TCG AAG GAC AGC GCG TAG CTC AAG GCC TTG
148 L P N H L F V N L S F L S R I E F R N

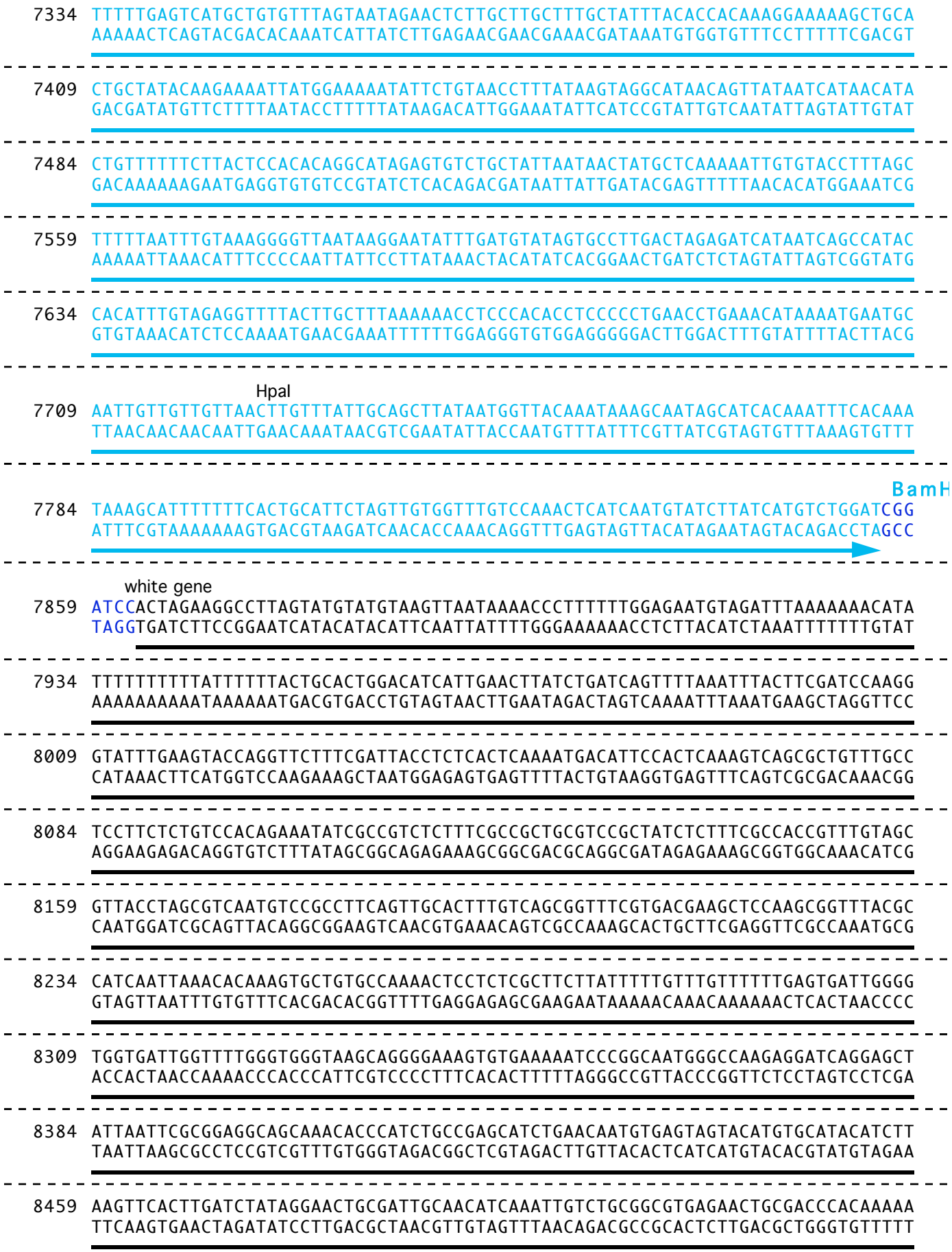
```
3954 AAT CGA TTG CGC CAG GTG CAG CTG CAC GTC TTC GCT GGC ACA ATG GCG CTG AGC GCC
     TTA GCT AAC GCG GTC CAC GTC GAC GTG CAG AAG CGA CCG TGT TAC CGC GAC TCG CGG

 167  N  R  L  R  Q  V  Q  L  H  V  F  A  G  T  M  A  L  S  A

4011 ATT TCG CTG GAA CAG AAC CGC CTC TCA CAT CTG CAC AAG GAG ACA TTC AAG GAT CTG
     TAA AGC GAC CTT GTC TTG GCG GAG AGT GTA GAC GTG TTC CTC TGT AAG TTC CTA GAC

 186  I  S  L  E  Q  N  R  L  S  H  L  H  K  E  T  F  K  D  L

                                                  C FLANK
                                                  oligo594
4068 CAG AAG CTG ATG CAT CTA TCG CTG CAG GGT AAC GCA TGG AAC TGC AGC TGC GAG CTG
     GTC TTC GAC TAC GTA GAT AGC GAC GTC CCA TTG CGT ACC TTG ACG TCG ACG CTC GAC

 205  Q  K  L  M  H  L  S  L  Q  G  N  A  W  N  C  S  C  E  L

4125 CAG GAC TTT CGC GAC TTT GCG ATC AGC AAA CGG CTC TAC ACA CCG CCC ACC GAT TGC
     GTC CTG AAA GCG CTG AAA CGC TAG TCG TTT GCC GAG ATG TGT GGC GGG TGG CTA ACG

 224  Q  D  F  R  D  F  A  I  S  K  R  L  Y  T  P  P  T  D  C

4182 CAG GAG CCG CCA CAG CTG CGC GGC AAG CTG TGG AGC GAG GTG CCA TCG GAG AAC TTC
     GTC CTC GGC GGT GTC GAC GCG CCG TTC GAC ACC TCG CTC CAC GGT AGC CTC TTG AAG

 243  Q  E  P  P  Q  L  R  G  K  L  W  S  E  V  P  S  E  N  F

                  Ig
4239 GCC TGC CGG CCG CGC ATT TTG GGT TCC GTG CGC TCC TTC ATC GAG GCC AAT CAC GAC
     CGG ACG GCC GGC GCG TAA AAC CCA AGG CAC GCG AGG AAG TAG CTC CGG TTA GTG CTG

 262  A  C  R  P  R  I  L  G  S  V  R  S  F  I  E  A  N  H  D

4296 AAT ATC TCG CTA CCC TGC CGC ATT GTC GGC AGT CCG CGT CCC AAT GTC ACC TGG GTG
     TTA TAG AGC GAT GGG ACG GCG TAA CAG CCG TCA GGC GCA GGG TTA CAG TGG ACC CAC

 281  N  I  S  L  P  C  R  I  V  G  S  P  R  P  N  V  T  W  V

4353 TAC AAC AAG CGG CCA TTG CAG CAG TAC GAC CCG CGT GTG CGT GTC CTC ACC TCC GTG
     ATG TTG TTC GCC GGT AAC GTC GTC ATG CTG GGC GCA CAC GCA CAG GAG TGG AGG CAC

 300  Y  N  K  R  P  L  Q  Q  Y  D  P  R  V  R  V  L  T  S  V

4410 GAA CAG ATG CCG GAG CAG CCC TCC CAG GTG CTC ACC TCG GAG CTG CGC ATC GTG GGC
     CTT GTC TAC GGC CTC GTC GGG AGG GTC CAC GAG TGG AGC CTC GAC GCG TAG CAC CCG

 319  E  Q  M  P  E  Q  P  S  Q  V  L  T  S  E  L  R  I  V  G

4467 GTA CGG GCC TCC GAC AAG GGT GCC TAC ACC TGT GTG GCG GAT AAC CGG GGC GGA CGG
     CAT GCC CGG AGG CTG TTC CCA CGG ATG TGG ACA CAC CGC CTA TTG GCC CCG CCT GCC

 338  V  R  A  S  D  K  G  A  Y  T  C  V  A  D  N  R  G  G  R

                                                  end Ig?
4524 GCG GAG GCC GAG TTC CAG CTG CTC GTG AGC GGT GAC TAT GCC GGC GCG GTA TCC GCC
     CGC CTC CGG CTC AAG GTC GAC GAG CAC TCG CCA CTG ATA CGG CCG CGC CAT AGG CGG

 357  A  E  A  E  F  Q  L  L  V  S  G  D  Y  A  G  A  V  S  A
```

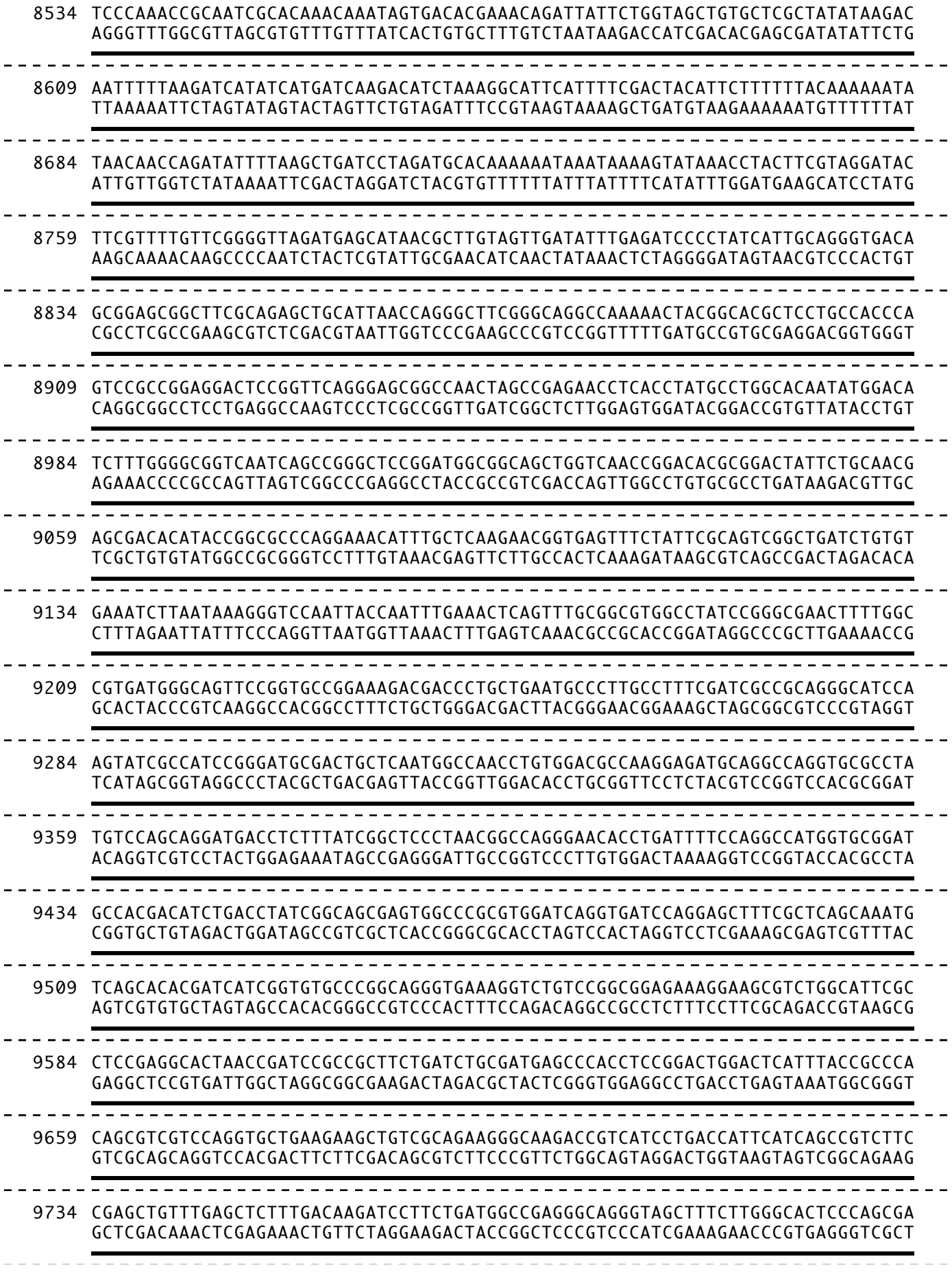

RNAi 27249 / 47770
W69
4581 TCC GAT GGC ATG GGC ATG GGC GCC ATT GGG GCA CCA ACC ATT GAT CCG CAA ACG AAC
AGG CTA CCG TAC CCG TAC CCG CGG TAA CCC CGT GGT TGG TAA CTA GGC GTT TGC TTG
376 S D G M G M G A I G A P T I D P Q T N
Tm
4638 ATG TTT CTC ATC ATC TGT CTA ATC ATT ACG ACG CTG CTG CTC CTG CTG CTC GTG GCG
TAC AAA GAG TAG TAG ACA GAT TAG TAA TGC TGC GAC GAC GAG GAC GAC GAG CAC CGC
395 M F L I I C L I I T T L L L L L V A
IC
oligo745 apK5_seq_ir
4695 GTG CTG ACG CTC TTC TGG TAC TGC CGT CGC ATC AAG ACC TAT CAA AAG GAC ACC ACC
CAC GAC TGC GAG AAG ACC ATG ACG GCA GCG TAG TTC TGG ATA GTT TTC CTG TGG TGG
414 V L T L F W Y C R R I K T Y Q K D T T
4752 ATG ATG AGC GGC GAC GGG CTG ATC TCT TCC AAG ATG GAC AAG ACG CAC AAC GGC TCC
TAC TAC TCG CCG CTG CCC GAC TAG AGA AGG TTC TAC CTG TTC TGC GTG TTG CCG AGG
433 M M S G D G L I S S K M D K T H N G S
IC1
oligo746 apK5_seq_int_3'
4809 ATG CTC GAG GGT TCC GTC ATC ATG GAG ATG CAG AAG AGC CTG CTC AAC GAG GTC AAT
TAC GAG CTC CCA AGG CAG TAG TAC CTC TAC GTC TTC TCG GAC GAG TTG CTC CAG TTA
452 M L E G S V I M E M Q K S L L N E V N
oligo240
4866 CCA GTC GAG AAG CCG CCA CGG CGC ACG GAC ATC GAG AGC GTG GAT GGT GGC GAT GAC
GGT CAG CTC TTC GGC GGT GCC GCG TGC CTG TAG CTC TCG CAC CTA CCA CCG CTA CTG
471 P V E K P P R R T D I E S V D G G D D
4923 GTG CTC GAG ATC AAG AAG ACG CTG CTC GAC GAC ACC GTC TAT GTG GCC AAT CAC TCG
CAC GAG CTC TAG TTC TTC TGC GAC GAG CTG CTG TGG CAG ATA CAC CGG TTA GTG AGC
490 V L E I K K T L L D D T V Y V A N H S
4980 CGC GAC GAA GAA GCC GTC TCA GTG GCC ATG TCG GAT ACG ACG ACC ACG CCC CGA TCT
GCG CTG CTT CTT CGG CAG AGT CAC CGG TAC AGC CTA TGC TGC TGG TGC GGG GCT AGA
509 R D E E A V S V A M S D T T T T P R S
IC2
5037 CGA CAC ACC TAC GTG GAT GAT GCG TAT GCC AAT AGC TTG CCA CCG GAT CTG CTG GCC
GCT GTG TGG ATG CAC CTA CTA CGC ATA CGG TTA TCG AAC GGT GGC CTA GAC GAC CGG
528 R H T Y V D D A Y A N S L P P D L L A
5094 TTT CCC GCT CGC GTG CCG CCC ACC TCG CCC TCG ATG CAA TCG TCG CAG TCG AAC ATA
AAA GGG CGA GCG CAC GGC GGG TGG AGC GGG AGC TAC GTT AGC AGC GTC AGC TTG TAT
547 F P A R V P P T S P S M Q S S Q S N I


5151 CCC GAC CAG GTG ATC TAC GGC ATC CGT TCG CCA CCG TCG CTA ACC AGT CCG GTC TAC
GGG CTG GTC CAC TAG ATG CCG TAG GCA AGC GGT GGC AGC GAT TGG TCA GGC CAG ATG
566 P D Q V I Y G I R S P P S L T S P V Y
5208 ACG CAT ATG ACG CCG CAC GGC ATC TAC GGC ACC AAG ACG ATG ACG GCT CCG CAT AAC
TGC GTA TAC TGC GGC GTG CCG TAG ATG CCG TGG TTC TGC TAC TGC CGA GGC GTA TTG
585 T H M T P H G I Y G T K T M T A P H N
5265 GGC TTT ATG ACG CTG CAG CAT CCC AAG TCG CGC AAC CTG GCG CTC ATT GCC ACC ACC
CCG AAA TAC TGC GAC GTC GTA GGG TTC AGC GCG TTG GAC CGC GAG TAA CGG TGG TGG
604 G F M T L Q H P K S R N L A L I A T T
5322 AAC AGC AGT CGC CAG CAC CAG CAC CAC CAT CAG CTG CAG CAG CAG CAG CAG CAC CAC
TTG TCG TCA GCG GTC GTG GTC GTG GTG GTA GTC GAC GTC GTC GTC GTC GTC GTG GTG
623 N S S R Q H Q H H H Q L Q Q Q Q Q H H
5379 CAC CAC CAC CAG CAG CAA CAA CAA CAG CAG CAG CAG CAG CAA CAT CCG CTG GCC ACC
GTG GTG GTG GTC GTC GTT GTT GTT GTC GTC GTC GTC GTC GTT GTA GGC GAC CGG TGG
642 H H H Q Q Q Q Q Q Q Q Q Q Q Q H P L A T
5436 ACA TCG CCC TTC CTG CCC GCA CCC GTC GTC TAT TCG CCG GCC ACG GGT GTG GTC ATG
TGT AGC GGG AAG GAC GGG CGT GGG CAG CAG ATA AGC GGC CGG TGC CCA CAC CAG TAC
661 T S P F L P A P V V Y S P A T G V V M
IC3
5493 AAA CAG GGA TAT ATG ACC ATT CCG CGC AAG CCG CGC GCT CCC AGC TGG GCG CCC AGT
TTT GTC CCT ATA TAC TGG TAA GGC GCG TTC GGC GCG CGA GGG TCG ACC CGC GGG TCA
680 K Q G Y M T I P R K P R A P S W A P S
5550 ACT TCC GGT GCC GCT GGC CAC GGA TCC ATT CAG CTA AGT GAA TTC CAG AGC CCC ACA
TGA AGG CCA CGG CGA CCG GTG CCT AGG TAA GTC GAT TCA CTT AAG GTC TCG GGG TGT
699 T S G A A G H G S I Q L S E F Q S P T
oligo241
IC4
5607 TCG CCG AAT CCC AGC GAG ACT GGC ACC GCC ACC ACC GCG GAA CTG CAG GCG GAG CCA
AGC GGC TTA GGG TCG CTC TGA CCG TGG CGG TGG TGG CGC CTT GAC GTC CGC CTC GGT
718 S P N P S E T G T A T T A E L Q A E P
5664 GTG TAC GAC AAC TTG GGA TTG CGA ACC ACT GCC GGC GGC AAC TCC ACC CTC AAT CTG
CAC ATG CTG TTG AAC CCT AAC GCT TGG TGA CGG CCG CCG TTG AGG TGG GAG TTA GAC
737 V Y D N L G L R T T A G G N S T L N L
5721 ACC AAG ATC GCC GGC TCA CAG GGG GGC GCT GGT CAG CAG TAC TCG ATG CGG GAC CGA
TGG TTC TAG CGG CCG AGT GTC CCC CCG CGA CCA GTC GTC ATG AGC TAC GCC CTG GCT
756 T K I A G S Q G G A G Q Q Y S M R D R
5778 CCA CTT CCG GCC ACG CCC AGC CTG ACA TCG GTG TCC TCG GCG ACC AAT GCC AGT AAG
GGT GAA GGC CGG TGC GGG TCG GAC TGT AGC CAC AGG AGC CGC TGG TTA CGG TCA TTC
775 P L P A T P S L T S V S S A T N A S K
5835 ATT TAC GAG CCC ATA CAC GAG CTG ATT CAG CAG CAA CAG CAG TTG CAA CAA CAA CAA
TAA ATG CTC GGG TAT GTG CTC GAC TAA GTC GTC GTT GTC GTC AAC GTT GTT GTT GTT
794 I Y E P I H E L I Q Q Q Q Q L Q Q Q Q

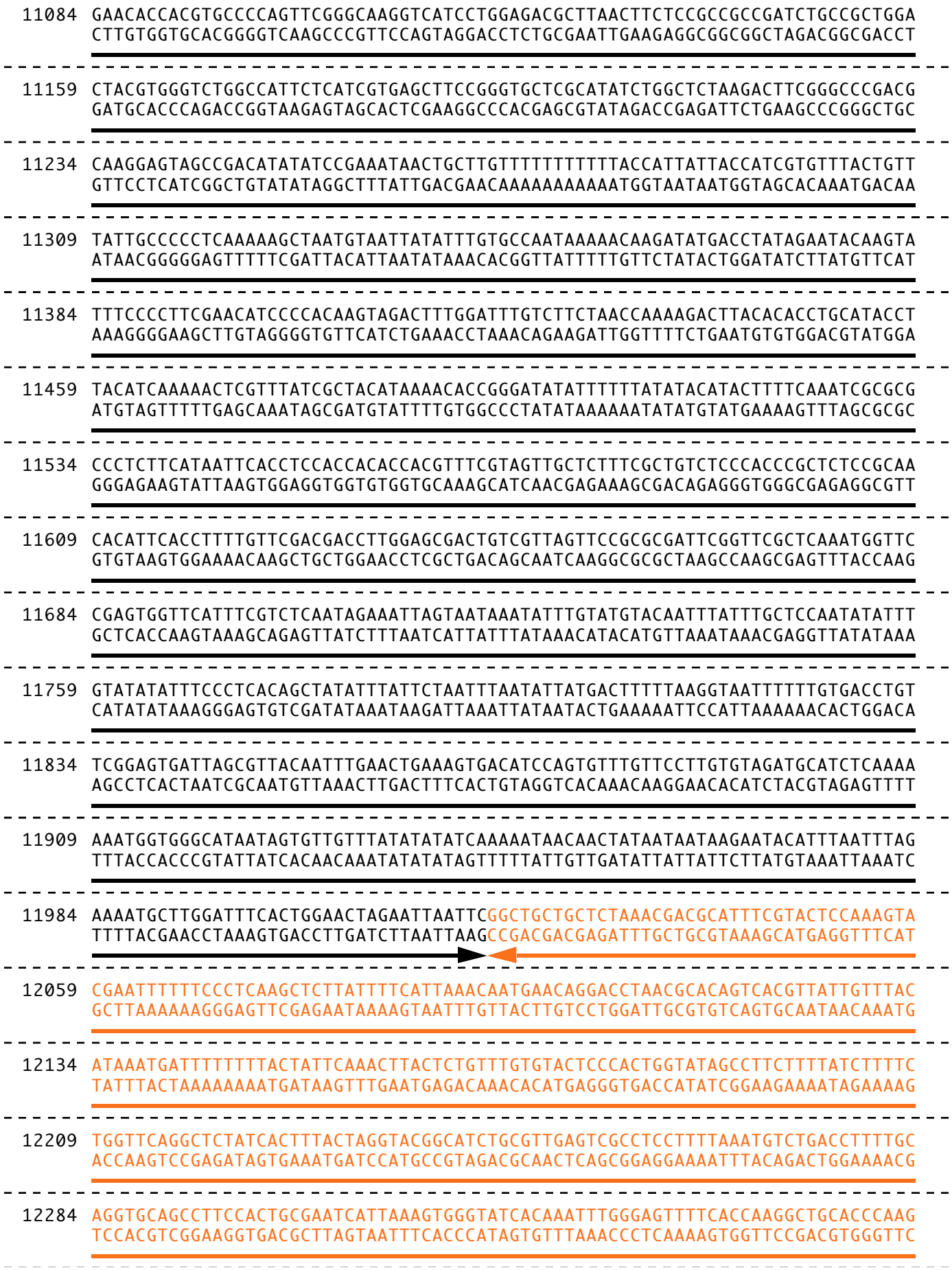

5892 CAG CAG CAG CAG CAG CGA CTG GGC TCC ATG GAC ACG GAA CCC CTG TAC GGA GTT CGG
GTC GTC GTC GTC GTC GCT GAC CCG AGG TAC CTG TGC CTT GGG GAC ATG CCT CAA GCC
813 Q Q Q Q Q R L G S M D T E P L Y G V R
5949 CAA CAG GGG ATC ACG ATA CTG CCC GGC TCG AGC ATT AGC GGT GCC GGA CTG GGC CAC
GTT GTC CCC TAG TGC TAT GAC GGG CCG AGC TCG TAA TCG CCA CGG CCT GAC CCG GTG
832 Q Q G I T I L P G S S I S G A G L G H
6006 GCC GCC TAC CTT TCA CCC GGC TCG GGT GCC GCC GTA TCG CCA AGC CAC GCC AGC AGC
CGG CGG ATG GAA AGT GGG CCG AGC CCA CGG CGG CAT AGC GGT TCG GTG CGG TCG TCG
851 A A Y L S P G S G A A V S P S H A S S
IC5
6063 AGC GGT GAC TCT CCG AAG GCC GCC AAG ATC CCA CCA CGC CCA CCA CCG AAG CCC AAG
TCG CCA CTG AGA GGC TTC CGG CGG TTC TAG GGT GGT GCG GGT GGT GGC TTC GGG TTC
870 S G D S P K A A K I P P R P P P K P K
6120 AAG AAG ATG TCC GTG ACG ACG ACG CGC AGC GGC CAG GGC AGC ACC AGC CAG CTC TTC
TTC TTC TAC AGG CAC TGC TGC TGC GCG TCG CCG GTC CCG TCG TGG TCG GTC GAG AAG
889 K K M S V T T T R S G Q G S T S Q L F
IC6 (PDZ)
attB2
6177 GAC GAC GAG GGC GAG GAT GGC ACC GAG GTC GAC CCA GCT TTC TTG TAC AAA
CTG CTG CTC CCG CTC CTA CCG TGG CTC CAG CTG GGT CGA AAG AAC ATG TTT
908 D D E G E D G T E V D P A F L Y K
EGFPN1 polylinker
KpnI
oligo 445 T7
Oligo 453
Oligo 416
EGFPN1
6228 GTG GTG GTA CCG CGG GCC CGG GAT CCA CCG GTC GCC ACC ATG GTG AGC AAG GGC
CAC CAC CAT GGC GCC CGG GCC CTA GGT GGC CAG CGG TGG TAC CAC TCG TTC CCG
925 V V V P R A R D P P V A T M V S K G
oligo744 apK5_seq_3'
6282 GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC GTA AAC
CTC CTC GAC AAG TGG CCC CAC CAC GGG TAG GAC CAG CTC GAC CTG CCG CTG CAT TTG
943 E E L F T G V V P I L V E L D G D V N
#54
6339 GGC CAC AAG TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG
CCG GTG TTC AAG TCG CAC AGG CCG CTC CCG CTC CCG CTA CGG TGG ATG CCG TTC GAC
962 G H K F S V S G E G E G D A T Y G K L
6396 ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG
TGG GAC TTC AAG TAG ACG TGG TGG CCG TTC GAC GGG CAC GGG ACC GGG TGG GAG CAC
981 T L K F I C T T G K L P V P W P T L V

```
  6453 ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG CAG
       TGG TGG GAC TGG ATG CCG CAC GTC ACG AAG TCG GCG ATG GGG CTG GTG TAC TTC GTC

 1000  T   T   L   T   Y   G   V   Q   C   F   S   R   Y   P   D   H   M   K   Q

  6510 CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG CGC ACC ATC TTC
       GTG CTG AAG AAG TTC AGG CGG TAC GGG CTT CCG ATG CAG GTC CTC GCG TGG TAG AAG

 1019  H   D   F   F   K   S   A   M   P   E   G   Y   V   Q   E   R   T   I   F

  6567 TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC
       AAG TTC CTG CTG CCG TTG ATG TTC TGG GCG CGG CTC CAC TTC AAG CTC CCG CTG TGG

 1038  F   K   D   D   G   N   Y   K   T   R   A   E   V   K   F   E   G   D   T

  6624 CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG
       GAC CAC TTG GCG TAG CTC GAC TTC CCG TAG CTG AAG TTC CTC CTG CCG TTG TAG GAC

 1057  L   V   N   R   I   E   L   K   G   I   D   F   K   E   D   G   N   I   L

  6681 GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG
       CCC GTG TTC GAC CTC ATG TTG ATG TTG TCG GTG TTG CAG ATA TAG TAC CGG CTG TTC

 1076  G   H   K   L   E   Y   N   Y   N   S   H   N   V   Y   I   M   A   D   K

  6738 CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC
       GTC TTC TTG CCG TAG TTC CAC TTG AAG TTC TAG GCG GTG TTG TAG CTC CTG CCG TCG

 1095  Q   K   N   G   I   K   V   N   F   K   I   R   H   N   I   E   D   G   S

  6795 GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC CCC GTG CTG
       CAC GTC GAG CGG CTG GTG ATG GTC GTC TTG TGG GGG TAG CCG CTG CCG GGG CAC GAC

 1114  V   Q   L   A   D   H   Y   Q   Q   N   T   P   I   G   D   G   P   V   L

  6852 CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG
       GAC GGG CTG TTG GTG ATG GAC TCG TGG GTC AGG CGG GAC TCG TTT CTG GGG TTG CTC

 1133  L   P   D   N   H   Y   L   S   T   Q   S   A   L   S   K   D   P   N   E

  6909 AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC
       TTC GCG CTA GTG TAC CAG GAC GAC CTC AAG CAC TGG CGG CGG CCC TAG TGA GAG CCG

 1152  K   R   D   H   M   V   L   L   E   F   V   T   A   A   G   I   T   L   G

                                      NotI          XbaI SV40 Poly A
  6966 ATG GAC GAG CTG TAC AAG TAA AGCGGCCGCGACTCTAGAGGATCTTTGTGAAGGAACCTTACTTCTGT
       TAC CTG CTC GAC ATG TTC ATT TCGCCGGCGCTGAGATCTCCTAGAAACACTTCCTTGGAATGAAGACA

 1171  M   D   E   L   Y   K   •

  7034 GGTGTGACATAATTGGACAAACTACCTACAGAGATTTAAAGCTCTAAGGTAAATATAAAATTTTTAAGTGTATAA
       CCACACTGTATTAACCTGTTTGATGGATGTCTCTAAATTTCGAGATTCCATTTATATTTTAAAAATTCACATATT

  7109 TGTGTTAAACTACTGATTCTAATTGTTTGTGTATTTTAGATTCCAACCTATGGAACTGATGAATGGGAGCAGTGG
       ACACAATTTGATGACTAAGATTAACAAACACATAAAATCTAAGGTTGGATACCTTGACTACTTACCCTCGTCACC

  7184 TGGAATGCCTTTAATGAGGAAAACCTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGAC
       ACCTTACGGAAATTACTCCTTTTGGACAAAACGAGTCTTCTTTACGGTAGATCACTACTACTCCGATGACGACTG

  7259 TCTCAACATTCTACTCCTCCAAAAAAGAAGAGAAAGGTAGAAGACCCCAAGGACTTTCCTTCAGAATTGCTAAGT
       AGAGTTGTAAGATGAGGAGGTTTTTTCTTCTCTTTCCATCTTCTGGGGTTCCTGAAAGGAAGTCTTAACGATTCA
```


7334 TTTTTGAGTCATGCTGTGTTTAGTAATAGAACTCTTGCTTGCTTTGCTATTTACACCACAAAGGAAAAAGCTGCA
AAAAACTCAGTACGACACAAATCATTATCTTGAGAACGAACGAAACGATAAATGTGGTGTTTCCTTTTTCGACGT
7409 CTGCTATACAAGAAAATTATGGAAAAATATTCTGTAACCTTTATAAGTAGGCATAACAGTTATAATCATAACATA
GACGATATGTTCTTTTAATACCTTTTTATAAGACATTGGAAATATTCATCCGTATTGTCAATATTAGTATTGTAT
7484 CTGTTTTTTCTTACTCCACACAGGCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAATTGTGTACCTTTAGC
GACAAAAAAGAATGAGGTGTGTCCGTATCTCACAGACGATAATTATTGATACGAGTTTTTAACACATGGAAATCG
7559 TTTTTAATTTGTAAAGGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAGATCATAATCAGCCATAC
AAAAATTAAACATTTCCCCAATTATTCCTTATAAACTACATATCACGGAACTGATCTCTAGTATTAGTCGGTATG
7634 CACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGC
GTGTAAACATCTCCAAAATGAACGAAATTTTTTGGAGGGTGTGGAGGGGGACTTGGACTTTGTATTTTACTTACG
HpaI
7709 AATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAA
TTAACAACAACAATTGAACAAATAACGTCGAATATTACCAATGTTTATTTCGTTATCGTAGTGTTTAAAGTGTTT
BamH
7784 TAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGGATCGG
ATTTCGTAAAAAAAAGTGACGTAAGATCAACACCAAACAGGTTTGAGTAGTTACATAGAATAGTACAGACCTAGCC
white gene
7859 ATCCACTAGAAGGCCTTAGTATGTATGTAAGTTAATAAAACCCTTTTTTGGAGAATGTAGATTTAAAAAAACATA
TAGGTGATCTTCCGGAATCATACATACATTCAATTATTTTGGGAAAAAACCTCTTACATCTAAATTTTTTTGTAT
7934 TTTTTTTTTTATTTTTTACTGCACTGGACATCATTGAACTTATCTGATCAGTTTTAAATTTACTTCGATCCAAGG
AAAAAAAAAATAAAAAATGACGTGACCTGTAGTAACTTGAATAGACTAGTCAAAATTTAAATGAAGCTAGGTTCC
8009 GTATTTGAAGTACCAGGTTCTTTCGATTACCTCTCACTCAAAATGACATTCCACTCAAAGTCAGCGCTGTTTGCC
CATAAACTTCATGGTCCAAGAAAGCTAATGGAGAGTGAGTTTTACTGTAAGGTGAGTTTCAGTCGCGACAAACGG
8084 TCCTTCTCTGTCCACAGAAATATCGCCGTCTCTTTCGCCGCTGCGTCCGCTATCTCTTTCGCCACCGTTTGTAGC
AGGAAGAGACAGGTGTCTTTATAGCGGCAGAGAAAGCGGCGACGCAGGCGATAGAGAAAGCGGTGGCAAACATCG
8159 GTTACCTAGCGTCAATGTCCGCCTTCAGTTGCACTTTGTCAGCGGTTTCGTGACGAAGCTCCAAGCGGTTTACGC
CAATGGATCGCAGTTACAGGCGGAAGTCAACGTGAAACAGTCGCCAAAGCACTGCTTCGAGGTTCGCCAAATGCG
8234 CATCAATTAAACACAAAGTGCTGTGCCAAAACTCCTCTCGCTTCTTATTTTTGTTTGTTTTTTGAGTGATTGGGG
GTAGTTAATTTGTGTTTCACGACACGGTTTTGAGGAGAGCGAAGAATAAAAACAAACAAAAAACTCACTAACCCC
8309 TGGTGATTGGTTTTGGGTGGGTAAGCAGGGGAAAGTGTGAAAAATCCCGGCAATGGGCCAAGAGGATCAGGAGCT
ACCACTAACCAAAACCCACCCATTCGTCCCCTTTCACACTTTTTAGGGCCGTTACCCGGTTCTCCTAGTCCTCGA
8384 ATTAATTCGCGGAGGCAGCAAACACCCATCTGCCGAGCATCTGAACAATGTGAGTAGTACATGTGCATACATCTT
TAATTAAGCGCCTCCGTCGTTTGTGGGTAGACGGCTCGTAGACTTGTTACACTCATCATGTACACGTATGTAGAA
8459 AAGTTCACTTGATCTATAGGAACTGCGATTGCAACATCAAATTGTCTGCGGCGTGAGAACTGCGACCCACAAAAA
TTCAAGTGAACTAGATATCCTTGACGCTAACGTTGTAGTTTAACAGACGCCGCACTCTTGACGCTGGGTGTTTTT


8534 TCCCAAACCGCAATCGCACAAACAAATAGTGACACGAAACAGATTATTCTGGTAGCTGTGCTCGCTATATAAGAC
AGGGTTTGGCGTTAGCGTGTTTGTTTATCACTGTGCTTTGTCTAATAAGACCATCGACACGAGCGATATATTCTG
8609 AATTTTTAAGATCATATCATGATCAAGACATCTAAAGGCATTCATTTTCGACTACATTCTTTTTTACAAAAAATA
TTAAAAATTCTAGTATAGTACTAGTTCTGTAGATTTCCGTAAGTAAAAGCTGATGTAAGAAAAAATGTTTTTTAT
8684 TAACAACCAGATATTTTAAGCTGATCCTAGATGCACAAAAAATAAATAAAAGTATAAACCTACTTCGTAGGATAC
ATTGTTGGTCTATAAAATTCGACTAGGATCTACGTGTTTTTTATTTATTTTCATATTTGGATGAAGCATCCTATG
8759 TTCGTTTTGTTCGGGGTTAGATGAGCATAACGCTTGTAGTTGATATTTGAGATCCCCTATCATTGCAGGGTGACA
AAGCAAAACAAGCCCCAATCTACTCGTATTGCGAACATCAACTATAAACTCTAGGGGATAGTAACGTCCCACTGT
8834 GCGGAGCGGCTTCGCAGAGCTGCATTAACCAGGGCTTCGGGCAGGCCAAAAACTACGGCACGCTCCTGCCACCCA
CGCCTCGCCGAAGCGTCTCGACGTAATTGGTCCCGAAGCCCGTCCGGTTTTTGATGCCGTGCGAGGACGGTGGGT
8909 GTCCGCCGGAGGACTCCGGTTCAGGGAGCGGCCAACTAGCCGAGAACCTCACCTATGCCTGGCACAATATGGACA
CAGGCGGCCTCCTGAGGCCAAGTCCCTCGCCGGTTGATCGGCTCTTGGAGTGGATACGGACCGTGTTATACCTGT
8984 TCTTTGGGGCGGTCAATCAGCCGGGCTCCGGATGGCGGCAGCTGGTCAACCGGACACGCGGACTATTCTGCAACG
AGAAACCCCGCCAGTTAGTCGGCCCGAGGCCTACCGCCGTCGACCAGTTGGCCTGTGCGCCTGATAAGACGTTGC
9059 AGCGACACATACCGGCGCCCAGGAAACATTTGCTCAAGAACGGTGAGTTTCTATTCGCAGTCGGCTGATCTGTGT
TCGCTGTGTATGGCCGCGGGTCCTTTGTAAACGAGTTCTTGCCACTCAAAGATAAGCGTCAGCCGACTAGACACA
9134 GAAATCTTAATAAAGGGTCCAATTACCAATTTGAAACTCAGTTTGCGGCGTGGCCTATCCGGGCGAACTTTTGGC
CTTTAGAATTATTTCCCAGGTTAATGGTTAAACTTTGAGTCAAACGCCGCACCGGATAGGCCCGCTTGAAAACCG
9209 CGTGATGGGCAGTTCCGGTGCCGGAAAGACGACCCTGCTGAATGCCCTTGCCTTTCGATCGCCGCAGGGCATCCA
GCACTACCCGTCAAGGCCACGGCCTTTCTGCTGGGACGACTTACGGGAACGGAAAGCTAGCGGCGTCCCGTAGGT
9284 AGTATCGCCATCCGGGATGCGACTGCTCAATGGCCAACCTGTGGACGCCAAGGAGATGCAGGCCAGGTGCGCCTA
TCATAGCGGTAGGCCCTACGCTGACGAGTTACCGGTTGGACACCTGCGGTTCCTCTACGTCCGGTCCACGCGGAT
9359 TGTCCAGCAGGATGACCTCTTTATCGGCTCCCTAACGGCCAGGGAACACCTGATTTTCCAGGCCATGGTGCGGAT
ACAGGTCGTCCTACTGGAGAAATAGCCGAGGGATTGCCGGTCCCTTGTGGACTAAAAGGTCCGGTACCACGCCTA
9434 GCCACGACATCTGACCTATCGGCAGCGAGTGGCCCGCGTGGATCAGGTGATCCAGGAGCTTTCGCTCAGCAAATG
CGGTGCTGTAGACTGGATAGCCGTCGCTCACCGGGCGCACCTAGTCCACTAGGTCCTCGAAAGCGAGTCGTTTAC
9509 TCAGCACACGATCATCGGTGTGCCCGGCAGGGTGAAAGGTCTGTCCGGCGGAGAAAGGAAGCGTCTGGCATTCGC
AGTCGTGTGCTAGTAGCCACACGGGCCGTCCCACTTTCCAGACAGGCCGCCTCTTTCCTTCGCAGACCGTAAGCG
9584 CTCCGAGGCACTAACCGATCCGCCGCTTCTGATCTGCGATGAGCCCACCTCCGGACTGGACTCATTTACCGCCCA
GAGGCTCCGTGATTGGCTAGGCGGCGAAGACTAGACGCTACTCGGGTGGAGGCCTGACCTGAGTAAATGGCGGGT
9659 CAGCGTCGTCCAGGTGCTGAAGAAGCTGTCGCAGAAGGGCAAGACCGTCATCCTGACCATTCATCAGCCGTCTTC
GTCGCAGCAGGTCCACGACTTCTTCGACAGCGTCTTCCCGTTCTGGCAGTAGGACTGGTAAGTAGTCGGCAGAAG
9734 CGAGCTGTTTGAGCTCTTTGACAAGATCCTTCTGATGGCCGAGGGCAGGGTAGCTTTCTTGGGCACTCCCAGCGA
GCTCGACAAACTCGAGAAACTGTTCTAGGAAGACTACCGGCTCCCGTCCCATCGAAAGAACCCGTGAGGGTCGCT

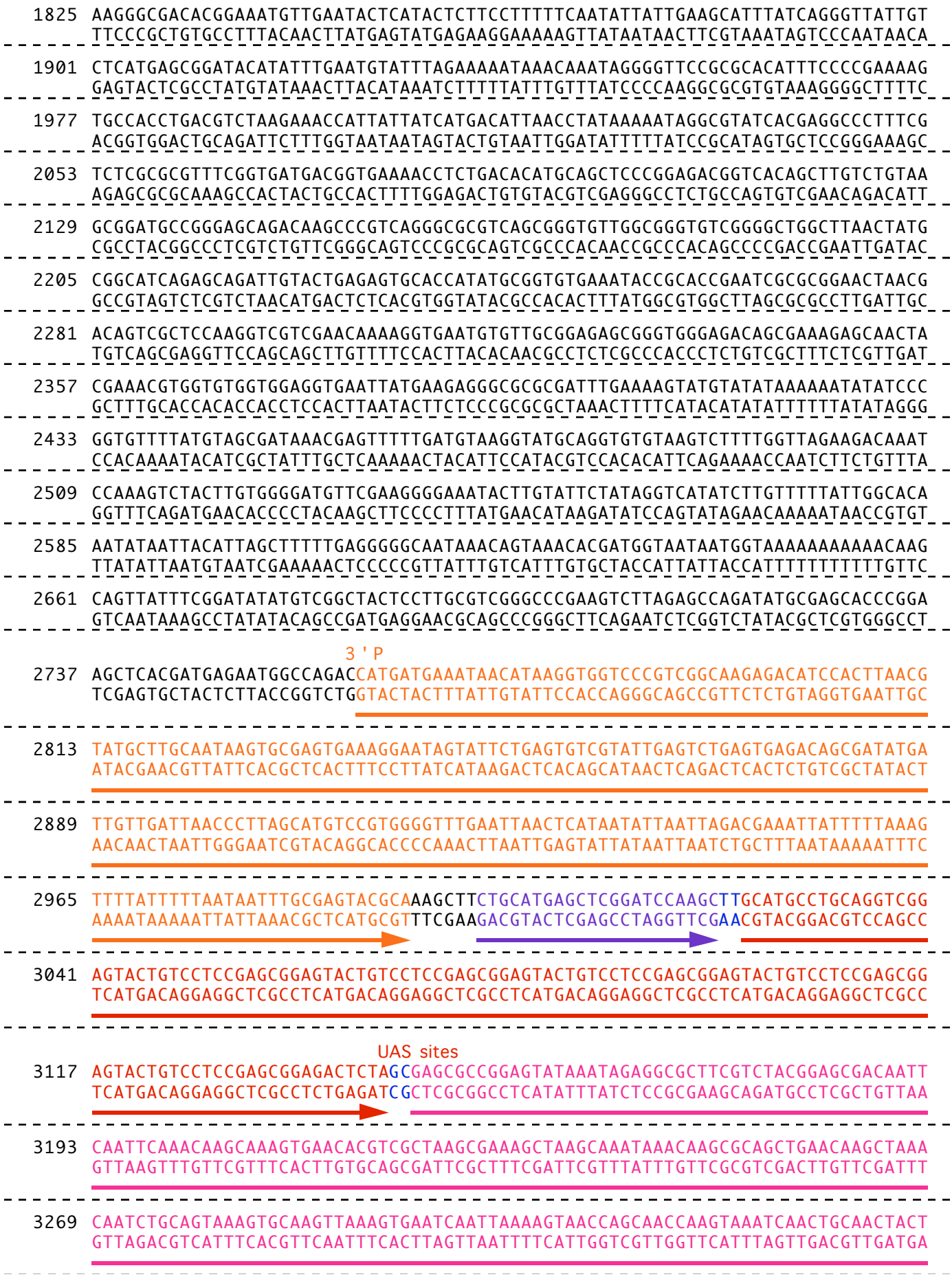
```
 9809 AGCCGTCGACTTCTTTTCCTAGTGAGTTCGATGTGTTTATTAAGGGTATCTAGCATTACATTACATCTCAACTCC
      TCGGCAGCTGAAGAAAAGGATCACTCAAGCTACACAAATAATTCCCATAGATCGTAATGTAATGTAGAGTTGAGG

 9884 TATCCAGCGTGGGTGCCCAGTGTCCTACCAACTACAATCCGGCGGACTTTTACGTACAGGTGTTGGCCGTTGTGC
      ATAGGTCGCACCCACGGGTCACAGGATGGTTGATGTTAGGCCGCCTGAAAATGCATGTCCACAACCGGCAACACG

 9959 CCGGACGGGAGATCGAGTCCCGTGATCGGATCGCCAAGATATGCGACAATTTTGCTATTAGCAAAGTAGCCCGGG
      GGCCTGCCCTCTAGCTCAGGGCACTAGCCTAGCGGTTCTATACGCTGTTAAAACGATAATCGTTTCATCGGGCCC

10034 ATATGGAGCAGTTGTTGGCCACCAAAAATTTGGAGAAGCCACTGGAGCAGCCGGAGAATGGGTACACCTACAAGG
      TATACCTCGTCAACAACCGGTGGTTTTTAAACCTCTTCGGTGACCTCGTCGGCCTCTTACCCATGTGGATGTTCC

10109 CCACCTGGTTCATGCAGTTCCGGGCGGTCCTGTGGCGATCCTGGCTGTCGGTGCTCAAGGAACCACTCCTCGTAA
      GGTGGACCAAGTACGTCAAGGCCCGCCAGGACACCGCTAGGACCGACAGCCACGAGTTCCTTGGTGAGGAGCATT

10184 AAGTGCGACTTATTCAGACAACGGTGAGTGGTTCCAGTGGAAACAAATGATATAACGCTTACAATTCTTGGAAAC
      TTCACGCTGAATAAGTCTGTTGCCACTCACCAAGGTCACCTTTGTTTACTATATTGCGAATGTTAAGAACCTTTG

10259 AAATTCGCTAGATTTTAGTTAGAATTGCCTGATTCCACACCCTTCTTAGTTTTTTTCAATGAGATGTATAGTTTA
      TTTAAGCGATCTAAAATCAATCTTAACGGACTAAGGTGTGGGAAGAATCAAAAAAAGTTACTCTACATATCAAAT

10334 TAGTTTTGCAGAAAATAAATAAATTTCATTTAACTCGCGAACATGTTGAAGATATGAATATTAATGAGATGCGAG
      ATCAAAACGTCTTTTATTTATTTAAAGTAAATTGAGCGCTTGTACAACTTCTATACTTATAATTACTCTACGCTC

10409 TAACATTTTAATTTGCAGATGGTTGCCATCTTGATTGGCCTCATCTTTTTGGGCCAACAACTCACGCAAGTGGGC
      ATTGTAAAATTAAACGTCTACCAACGGTAGAACTAACCGGAGTAGAAAAACCCGGTTGTTGAGTGCGTTCACCCG

10484 GTGATGAATATCAACGGAGCCATCTTCCTCTTCCTGACCAACATGACCTTTCAAAACGTCTTTGCCACGATAAAT
      CACTACTTATAGTTGCCTCGGTAGAAGGAGAAGGACTGGTTGTACTGGAAAGTTTTGCAGAAACGGTGCTATTTA

10559 GTAAGTCTTGTTTAGAATACATTTGCATATTAATAATTTACTAACTTTCTAATGAATCGATTCGATTTAGGTGTT
      CATTCAGAACAAATCTTATGTAAACGTATAATTATTAAATGATTGAAAGATTACTTAGCTAAGCTAAATCCACAA

10634 CACCTCAGAGCTGCCAGTTTTTATGAGGGAGGCCCGAAGTCGACTTTATCGCTGTGACACATACTTTCTGGGCAA
      GTGGAGTCTCGACGGTCAAAAATACTCCCTCCGGGCTTCAGCTGAAATAGCGACACTGTGTATGAAAGACCCGTT

10709 AACGATTGCCGAATTACCGCTTTTTCTCACAGTGCCACTGGTCTTCACGGCGATTGCCTATCCGATGATCGGACT
      TTGCTAACGGCTTAATGGCGAAAAAGAGTGTCACGGTGACCAGAAGTGCCGCTAACGGATAGGCTACTAGCCTGA

10784 GCGGGCCGGAGTGCTGCACTTCTTCAACTGCCTGGCGCTGGTCACTCTGGTGGCCAATGTGTCAACGTCCTTCGG
      CGCCCGGCCTCACGACGTGAAGAAGTTGACGGACCGCGACCAGTGAGACCACCGGTTACACAGTTGCAGGAAGCC

10859 ATATCTAATATCCTGCGCCAGCTCCTCGACCTCGATGGCGCTGTCTGTGGGTCCGCCGGTTATCATACCATTCCT
      TATAGATTATAGGACGCGGTCGAGGAGCTGGAGCTACCGCGACAGACACCCAGGCGGCCAATAGTATGGTAAGGA

10934 GCTCTTTGGCGGCTTCTTCTTGAACTCGGGCTCGGTGCCAGTATACCTCAAATGGTTGTCGTACCTCTCATGGTT
      CGAGAAACCGCCGAAGAAGAACTTGAGCCCGAGCCACGGTCATATGGAGTTTACCAACAGCATGGAGAGTACCAA

11009 CCGTTACGCCAACGAGGGTCTGCTGATTAACCAATGGGCGGACGTGGAGCCGGGCGAAATTAGCTGCACATCGTC
      GGCAATGCGGTTGCTCCCAGACGACTAATTGGTTACCCGCCTGCACCTCGGCCCGCTTTAATCGACGTGTAGCAG
```

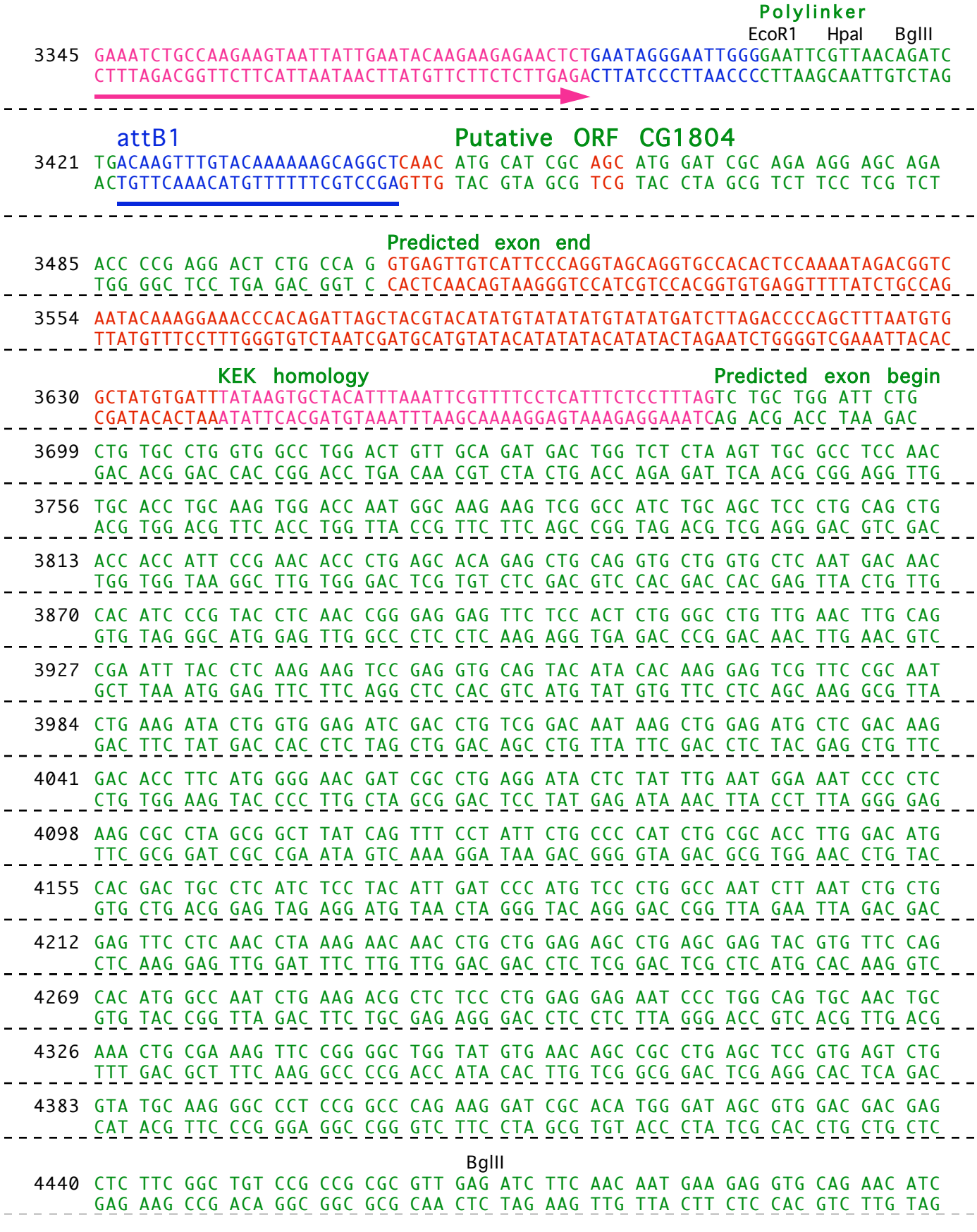

11084 GAACACCACGTGCCCCAGTTCGGGCAAGGTCATCCTGGAGACGCTTAACTTCTCCGCCGCCGATCTGCCGCTGGA
CTTGTGGTGCACGGGGTCAAGCCCGTTCCAGTAGGACCTCTGCGAATTGAAGAGGCGGCGGCTAGACGGCGACCT
11159 CTACGTGGGTCTGGCCATTCTCATCGTGAGCTTCCGGGTGCTCGCATATCTGGCTCTAAGACTTCGGGCCCGACG
GATGCACCCAGACCGGTAAGAGTAGCACTCGAAGGCCCACGAGCGTATAGACCGAGATTCTGAAGCCCGGGCTGC
11234 CAAGGAGTAGCCGACATATATCCGAAATAACTGCTTGTTTTTTTTTTTACCATTATTACCATCGTGTTTACTGTT
GTTCCTCATCGGCTGTATATAGGCTTTATTGACGAACAAAAAAAAAAAATGGTAATAATGGTAGCACAAATGACAA
11309 TATTGCCCCCTCAAAAAGCTAATGTAATTATATTTGTGCCAATAAAAACAAGATATGACCTATAGAATACAAGTA
ATAACGGGGGAGTTTTTCGATTACATTAATATAAACACGGTTATTTTTGTTCTATACTGGATATCTTATGTTCAT
11384 TTTCCCCTTCGAACATCCCCACAAGTAGACTTTGGATTTGTCTTCTAACCAAAAGACTTACACACCTGCATACCT
AAAGGGGAAGCTTGTAGGGGTGTTCATCTGAAACCTAAACAGAAGATTGGTTTTCTGAATGTGTGGACGTATGGA
11459 TACATCAAAAACTCGTTTATCGCTACATAAAACACCGGGATATATTTTTTATATACATACTTTTCAAATCGCGCG
ATGTAGTTTTTGAGCAAATAGCGATGTATTTTGTGGCCCTATATAAAAAATATATGTATGAAAAGTTTAGCGCGC
11534 CCCTCTTCATAATTCACCTCCACCACACCACGTTTCGTAGTTGCTCTTTCGCTGTCTCCCACCCGCTCTCCGCAA
GGGAGAAGTATTAAGTGGAGGTGGTGTGGTGCAAAGCATCAACGAGAAAGCGACAGAGGGTGGGCGAGAGGCGTT
11609 CACATTCACCTTTTGTTCGACGACCTTGGAGCGACTGTCGTTAGTTCCGCGCGATTCGGTTCGCTCAAATGGTTC
GTGTAAGTGGAAAACAAGCTGCTGGAACCTCGCTGACAGCAATCAAGGCGCGCTAAGCCAAGCGAGTTTACCAAG
11684 CGAGTGGTTCATTTCGTCTCAATAGAAATTAGTAATAAATATTTGTATGTACAATTTATTTGCTCCAATATATTT
GCTCACCAAGTAAAGCAGAGTTATCTTTAATCATTATTTATAAACATACATGTTAAATAAACGAGGTTATATAAA
11759 GTATATATTTCCCTCACAGCTATATTTATTCTAATTTAATATTATGACTTTTTAAGGTAATTTTTTGTGACCTGT
CATATATAAAGGGAGTGTCGATATAAATAAGATTAAATTATAATACTGAAAAATTCCATTAAAAAACACTGGACA
11834 TCGGAGTGATTAGCGTTACAATTTGAACTGAAAGTGACATCCAGTGTTTGTTCCTTGTGTAGATGCATCTCAAAA
AGCCTCACTAATCGCAATGTTAAACTTGACTTTCACTGTAGGTCACAAACAAGGAACACATCTACGTAGAGTTTT
11909 AAATGGTGGGCATAATAGTGTTGTTTATATATATCAAAAATAACAACTATAATAATAAGAATACATTTAATTTAG
TTTACCACCCGTATTATCACAACAAATATATATAGTTTTTATTGTTGATATTATTATTCTTATGTAAATTAAATC
11984 AAAATGCTTGGATTTCACTGGAACTAGAATTAATTCGGCTGCTGCTCTAAACGACGCATTTCGTACTCCAAAGTA
TTTTACGAACCTAAAGTGACCTTGATCTTAATTAAGCCGACGACGAGATTTGCTGCGTAAAGCATGAGGTTTCAT
12059 CGAATTTTTTCCCTCAAGCTCTTATTTTCATTAAACAATGAACAGGACCTAACGCACAGTCACGTTATTGTTTAC
GCTTAAAAAAGGGAGTTCGAGAATAAAAGTAATTTGTTACTTGTCCTGGATTGCGTGTCAGTGCAATAACAAATG
12134 ATAAATGATTTTTTTTACTATTCAAACTTACTCTGTTTGTGTACTCCCACTGGTATAGCCTTCTTTTATCTTTTC
TATTTACTAAAAAAAAATGATAAGTTTGAATGAGACAAACACATGAGGGTGACCATATCGGAAGAAAATAGAAAAG
12209 TGGTTCAGGCTCTATCACTTTACTAGGTACGGCATCTGCGTTGAGTCGCCTCCTTTTAAATGTCTGACCTTTTGC
ACCAAGTCCGAGATAGTGAAATGATCCATGCCGTAGACGCAACTCAGCGGAGGAAAATTTACAGACTGGAAAACG
12284 AGGTGCAGCCTTCCACTGCGAATCATTAAAGTGGGTATCACAAATTTGGGAGTTTTCACCAAGGCTGCACCCAAG
TCCACGTCGGAAGGTGACGCTTAGTAATTTCACCCATAGTGTTTAAACCCTCAAAAGTGGTTCCGACGTGGGTTC

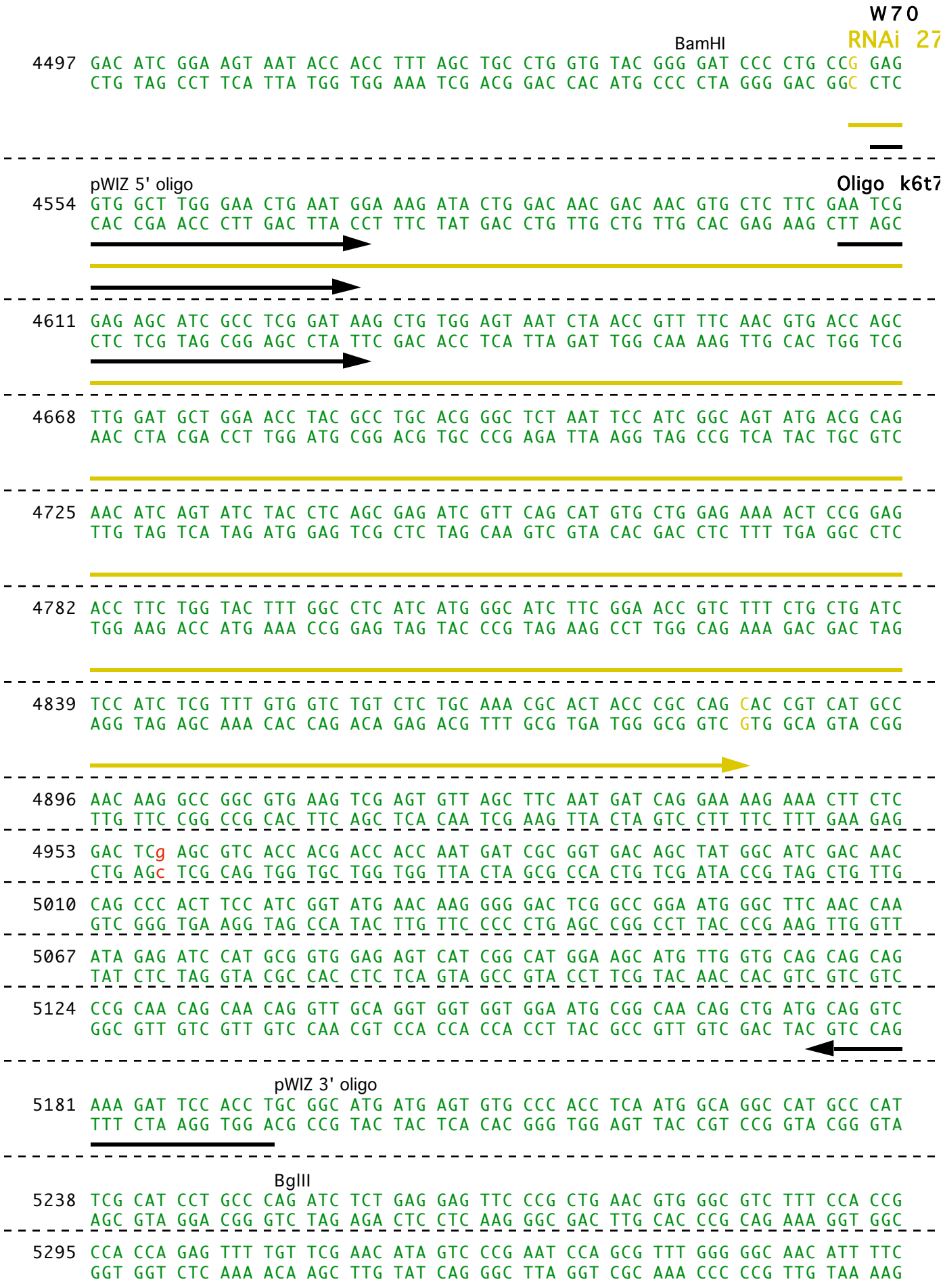

12359 GCTCTGCTCCCACAATTTTCTCTTAATAGCACACTTCGGCACGTGAATTAATTTTACTCCAGTCACAGCTTTGCA
CGAGACGAGGGTGTTAAAAGAGAATTATCGTGTGAAGCCGTGCACTTAATTAAAATGAGGTCAGTGTCGAAACGT
12434 GCAAAATTTGCAATATTTCATTTTTTTTTATTCCACGTAAGGGTTAATGTTTTCAAAAAAAAATTCGTCCGCACA
CGTTTTAAACGTTATAAAGTAAAAAAAAATAAGGTGCATTCCCAATTACAAAAGTTTTTTTTTAAGCAGGCGTGT
12509 CAACCTTTCCTCTCAACAAGCAAACGTGCACTGAATTTAAGTGTATACTTCGGTAAGCTTCGGCTATCGACGGGA
GTTGGAAAGGAGAGTTGTTCGTTTGCACGTGACTTAAATTCACATATGAAGCCATTCGAAGCCGATAGCTGCCCT
5' P
12584 CCACCTTATGTTATTTCATCATG
GGTGGAATACAATAAAGTAGTAC

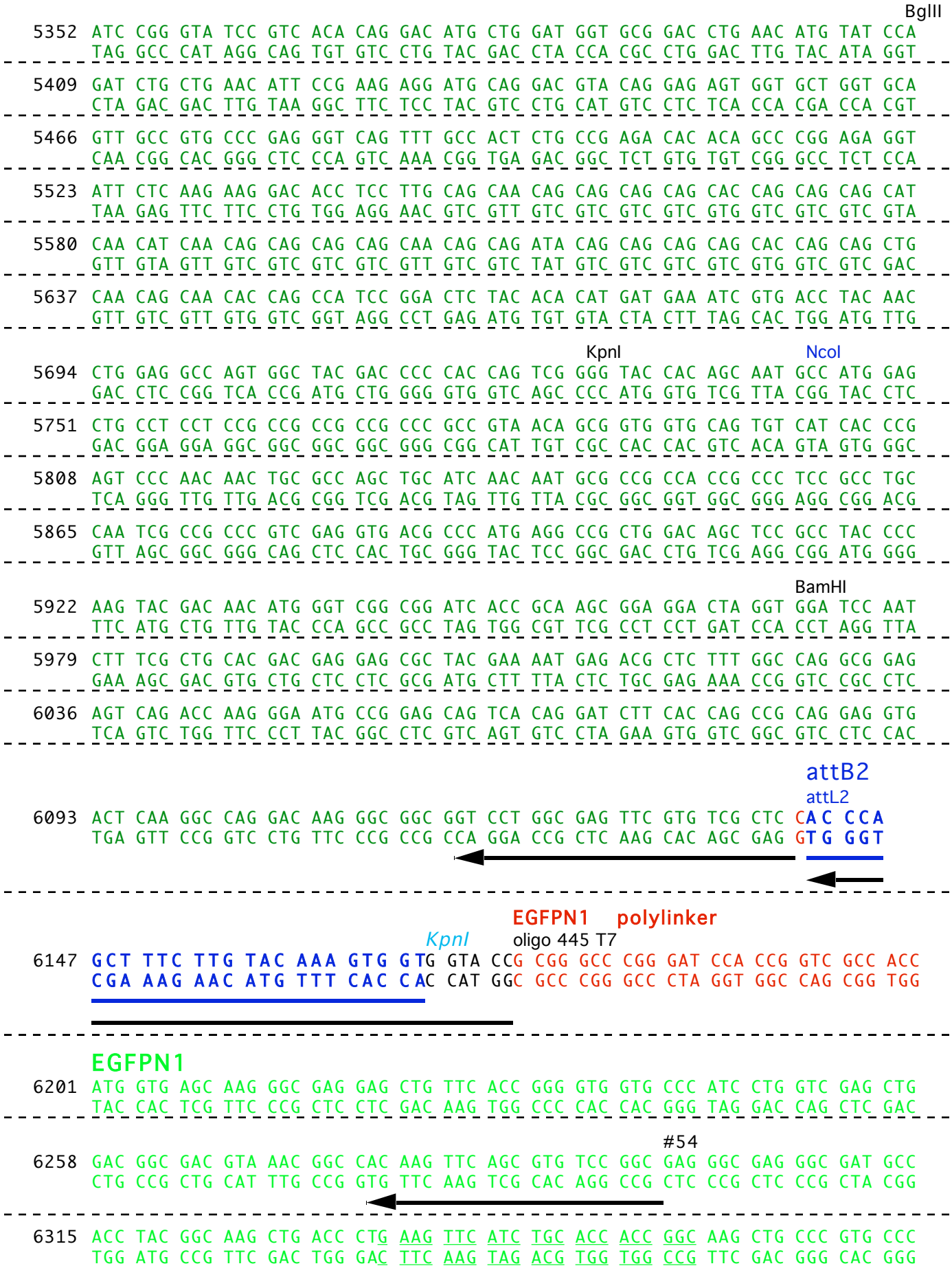
UAS Kek6 GFP

```
     1 GGCCAGACCCACGTAGTCCAGCGGCAGATCGGCGGCGGAGAAGTTAAGCGTCTCCAGGATGACCTTGCCCGAACTG
       CCGGTCTGGGTGCATCAGGTCGCCGTCTAGCCGCCGCCTCTTCAATTCGCAGAGGTCCTACTGGAACGGGCTTGAC

    77 GGGCACGTGGTGTTCGACGATGTGCAGCTAATTTCGCCCGGCTCCACGTCCGCCCATTGGTTAATCAGCAGACCCT
       CCCGTGCACCACAAGCTGCTACACGTCGATTAAAGCGGGCCGAGGTGCAGGCGGGTAACCAATTAGTCGTCTGGGA

   153 CGTTGGCGTAACGGAACCATGAGAGGTACGACAACCATTTGAGGTATACTGGCACCGAGCCCGAGTTCAAGAAGAA
       GCAACCGCATTGCCTTGGTACTCTCCATGCTGTTGGTAAACTCCATATGACCGTGGCTCGGGCTCAAGTTCTTCTT

   229 GGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCG
       CCGCAAAAAGGTATCCGAGGCGGGGGGACTGCTCGTAGTGTTTTTAGCTGCGAGTTCAGTCTCCACCGCTTTGGGC

   305 ACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTA
       TGTCCTGATATTTCTATGGTCCGCAAAGGGGGACCTTCGAGGGAGCACGCGAGAGGACAAGGCTGGGACGGCGAAT

   381 CCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTC
       GGCCTATGGACAGGCGGAAAGAGGGAAGCCCTTCGCACCGCGAAAGAGTTACGAGTGCGACATCCATAGAGTCAAG

   457 GGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGT
       CCACATCCAGCAAGCGAGGTTCGACCCGACACACGTGCTTGGGGGGCAAGTCGGGCTGGCGACGCGGAATAGGCCA

   533 AACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA
       TTGATAGCAGAACTCAGGTTGGGCCATTCTGTGCTGAATAGCGGTGACCGTCGTCGGTGACCATTGTCCTAATCGT

   609 GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATT
       CTCGCTCCATACATCCGCCACGATGTCTCAAGAACTTCACCACCGGATTGATGCCGATGTGATCTTCCTGTCATAA

   685 TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
       ACCATAGACGCGAGACGACTTCGGTCAATGGAAGCCTTTTTCTCAACCATCGAGAACTAGGCCGTTTGTTTGGTGG

   761 GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGA
       CGACCATCGCCACCAAAAAAACAAACGTTCGTCGTCTAATGCGCGTCTTTTTTTCCTAGAGTTCTTCTAGGAAACT

   837 TCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAG
       AGAAAAGATGCCCCAGACTGCGAGTCACCTTGCTTTTGAGTGCAATTCCCTAAAACCAGTACTCTAATAGTTTTTC

   913 GATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCT
       CTAGAAGTGGATCTAGGAAAATTTAATTTTTACTTCAAAATTTAGTTAGATTTCATATATACTCATTTGAACCAGA

   989 GACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGAC
       CTGTCAATGGTTACGAATTAGTCACTCCGTGGATAGAGTCGCTAGACAGATAAAGCAAGTAGGTATCAACGGACTG

  1065 TCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC
       AGGGGCAGCACATCTATTGATGCTATGCCCTCCCGAATGGTAGACCGGGGTCACGACGTTACTATGGCGCTCTGGG

  1141 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACT
       TGCGAGTGGCCGAGGTCTAAATAGTCGTTATTTGGTCGGTCGGCCTTCCCGGCTCGCGTCTTCACCAGGACGTTGA

  1217 TTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCA
       AATAGGCGGAGGTAGGTCAGATAATTAACAACGGCCCTTCGATCTCATTCATCAAGCGGTCAATTATCAAACGCGT

  1293 ACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCA
       TGCAACAACGGTAACGATGTCCGTAGCACCACAGTGCGAGCAGCAAACCATACCGAAGTAAGTCGAGGCCAAGGGT

  1369 ACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTC
       TGCTAGTTCCGCTCAATGTACTAGGGGGTACAACACGTTTTTTCGCCAATCGAGGAAGCCAGGAGGCTAGCAACAG

  1445 AGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCG
       TCTTCATTCAACCGGCGTCACAATAGTGAGTACCAATACCGTCGTGACGTATTAAGAGAATGACAGTACGGTAGGC

  1521 TAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTC
       ATTCTACGAAAAGACACTGACCACTCATGAGTTGGTTCAGTAAGACTCTTATCACATACGCCGCTGGCTCAACGAG

  1597 TTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCT
       AACGGGCCGCAGTTGTGCCCTATTATGGCGCGGTGTATCGTCTTGAAATTTTCACGAGTAGTAACCTTTTGCAAGA

  1673 TCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGAT
       AGCCCCGCTTTTGAGAGTTCCTAGAATGGCGACAACTCTAGGTCAAGCTACATTGGGTGAGCACGTGGGTTGACTA

  1749 CTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAAT
       GAAGTCGTAGAAAATGAAAGTGGTCGCAAAGACCCACTCGTTTTTGTCCTTCCGTTTTACGGCGTTTTTTCCCTTA
```

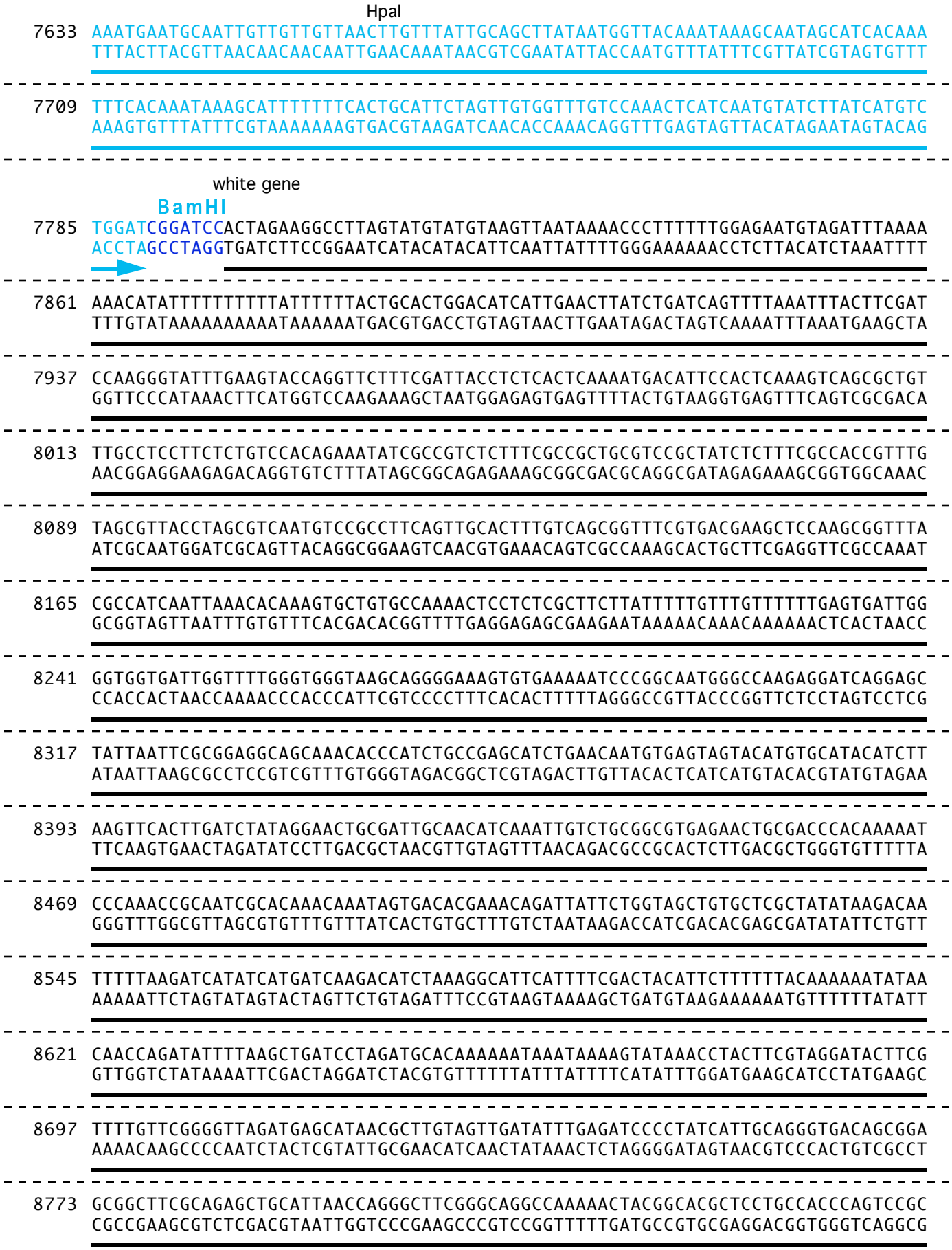

3 ' P
UAS sites


Polylinker
EcoR1 HpaI BglII
attB1
Putative ORF CG1804
Predicted exon end
KEK homology
Predicted exon begin
BglII

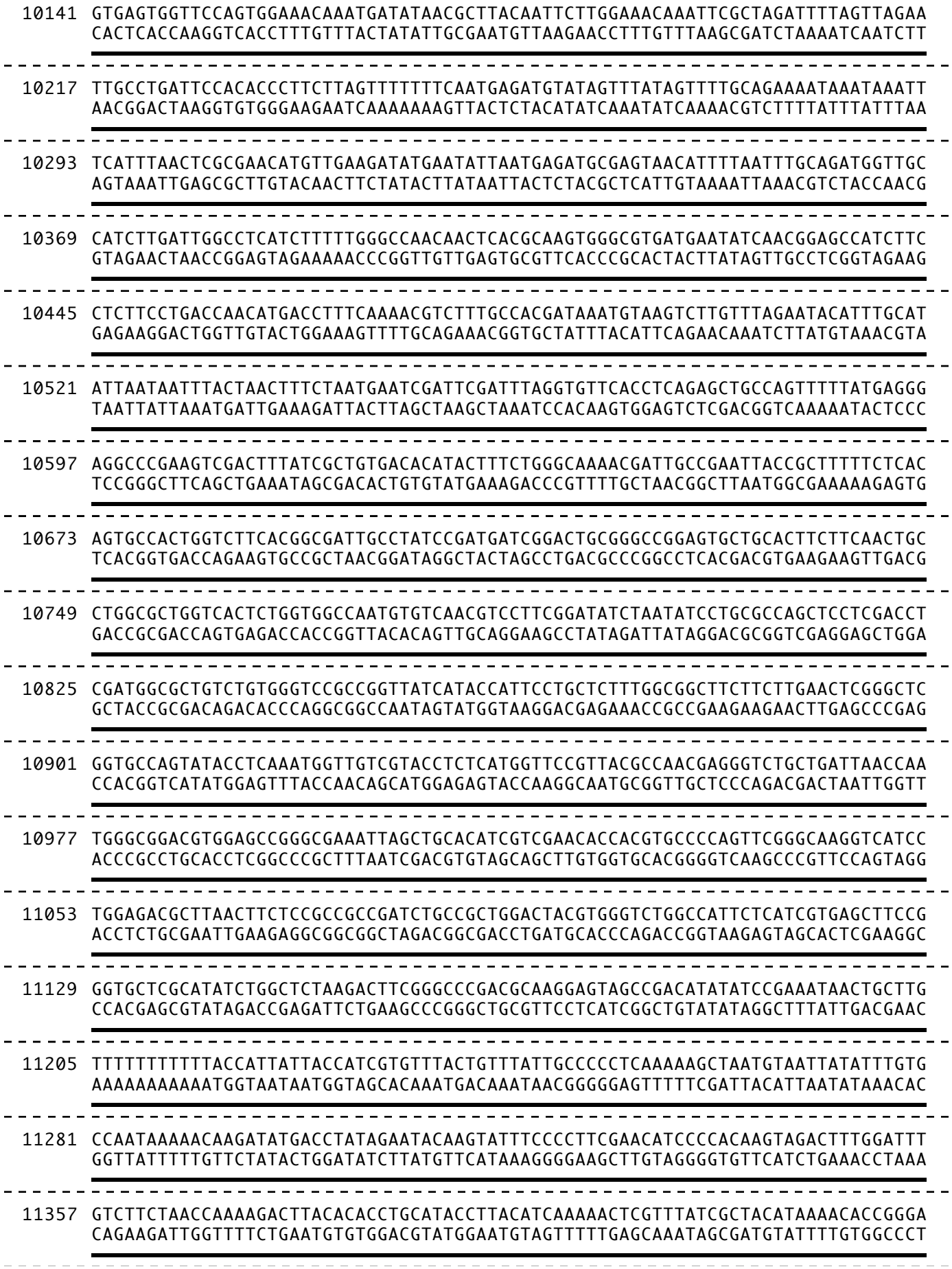

W70
RNAi 27
BamHI
4497 GAC ATC GGA AGT AAT ACC ACC TTT AGC TGC CTG GTG TAC GGG GAT CCC CTG CCG GAG
CTG TAG CCT TCA TTA TGG TGG AAA TCG ACG GAC CAC ATG CCC CTA GGG GAC GGC CTC
pWIZ 5' oligo
Oligo k6t7
4554 GTG GCT TGG GAA CTG AAT GGA AAG ATA CTG GAC AAC GAC AAC GTG CTC TTC GAA TCG
CAC CGA ACC CTT GAC TTA CCT TTC TAT GAC CTG TTG CTG TTG CAC GAG AAG CTT AGC
4611 GAG AGC ATC GCC TCG GAT AAG CTG TGG AGT AAT CTA ACC GTT TTC AAC GTG ACC AGC
CTC TCG TAG CGG AGC CTA TTC GAC ACC TCA TTA GAT TGG CAA AAG TTG CAC TGG TCG
4668 TTG GAT GCT GGA ACC TAC GCC TGC ACG GGC TCT AAT TCC ATC GGC AGT ATG ACG CAG
AAC CTA CGA CCT TGG ATG CGG ACG TGC CCG AGA TTA AGG TAG CCG TCA TAC TGC GTC
4725 AAC ATC AGT ATC TAC CTC AGC GAG ATC GTT CAG CAT GTG CTG GAG AAA ACT CCG GAG
TTG TAG TCA TAG ATG GAG TCG CTC TAG CAA GTC GTA CAC GAC CTC TTT TGA GGC CTC
4782 ACC TTC TGG TAC TTT GGC CTC ATC ATG GGC ATC TTC GGA ACC GTC TTT CTG CTG ATC
TGG AAG ACC ATG AAA CCG GAG TAG TAC CCG TAG AAG CCT TGG CAG AAA GAC GAC TAG
4839 TCC ATC TCG TTT GTG GTC TGT CTC TGC AAA CGC ACT ACC CGC CAG CAC CGT CAT GCC
AGG TAG AGC AAA CAC CAG ACA GAG ACG TTT GCG TGA TGG GCG GTC GTG GCA GTA CGG
4896 AAC AAG GCC GGC GTG AAG TCG AGT GTT AGC TTC AAT GAT CAG GAA AAG AAA CTT CTC
TTG TTC CGG CCG CAC TTC AGC TCA CAA TCG AAG TTA CTA GTC CTT TTC TTT GAA GAG
4953 GAC TCg AGC GTC ACC ACG ACC ACC AAT GAT CGC GGT GAC AGC TAT GGC ATC GAC AAC
CTG AGc TCG CAG TGG TGC TGG TGG TTA CTA GCG CCA CTG TCG ATA CCG TAG CTG TTG
5010 CAG CCC ACT TCC ATC GGT ATG AAC AAG GGG GAC TCG GCC GGA ATG GGC TTC AAC CAA
GTC GGG TGA AGG TAG CCA TAC TTG TTC CCC CTG AGC CGG CCT TAC CCG AAG TTG GTT
5067 ATA GAG ATC CAT GCG GTG GAG AGT CAT CGG CAT GGA AGC ATG TTG GTG CAG CAG CAG
TAT CTC TAG GTA CGC CAC CTC TCA GTA GCC GTA CCT TCG TAC AAC CAC GTC GTC GTC
5124 CCG CAA CAG CAA CAG GTT GCA GGT GGT GGT GGA ATG CGG CAA CAG CTG ATG CAG GTC
GGC GTT GTC GTT GTC CAA CGT CCA CCA CCA CCT TAC GCC GTT GTC GAC TAC GTC CAG
pWIZ 3' oligo
5181 AAA GAT TCC ACC TGC GGC ATG ATG AGT GTG CCC ACC TCA ATG GCA GGC CAT GCC CAT
TTT CTA AGG TGG ACG CCG TAC TAC TCA CAC GGG TGG AGT TAC CGT CCG GTA CGG GTA
BglII
5238 TCG CAT CCT GCC CAG ATC TCT GAG GAG TTC CCG CTG AAC GTG GGC GTC TTT CCA CCG
AGC GTA GGA CGG GTC TAG AGA CTC CTC AAG GGC GAC TTG CAC CCG CAG AAA GGT GGC
5295 CCA CCA GAG TTT TGT TCG AAC ATA GTC CCG AAT CCA GCG TTT GGG GGC AAC ATT TTC
GGT GGT CTC AAA ACA AGC TTG TAT CAG GGC TTA GGT CGC AAA CCC CCG TTG TAA AAG

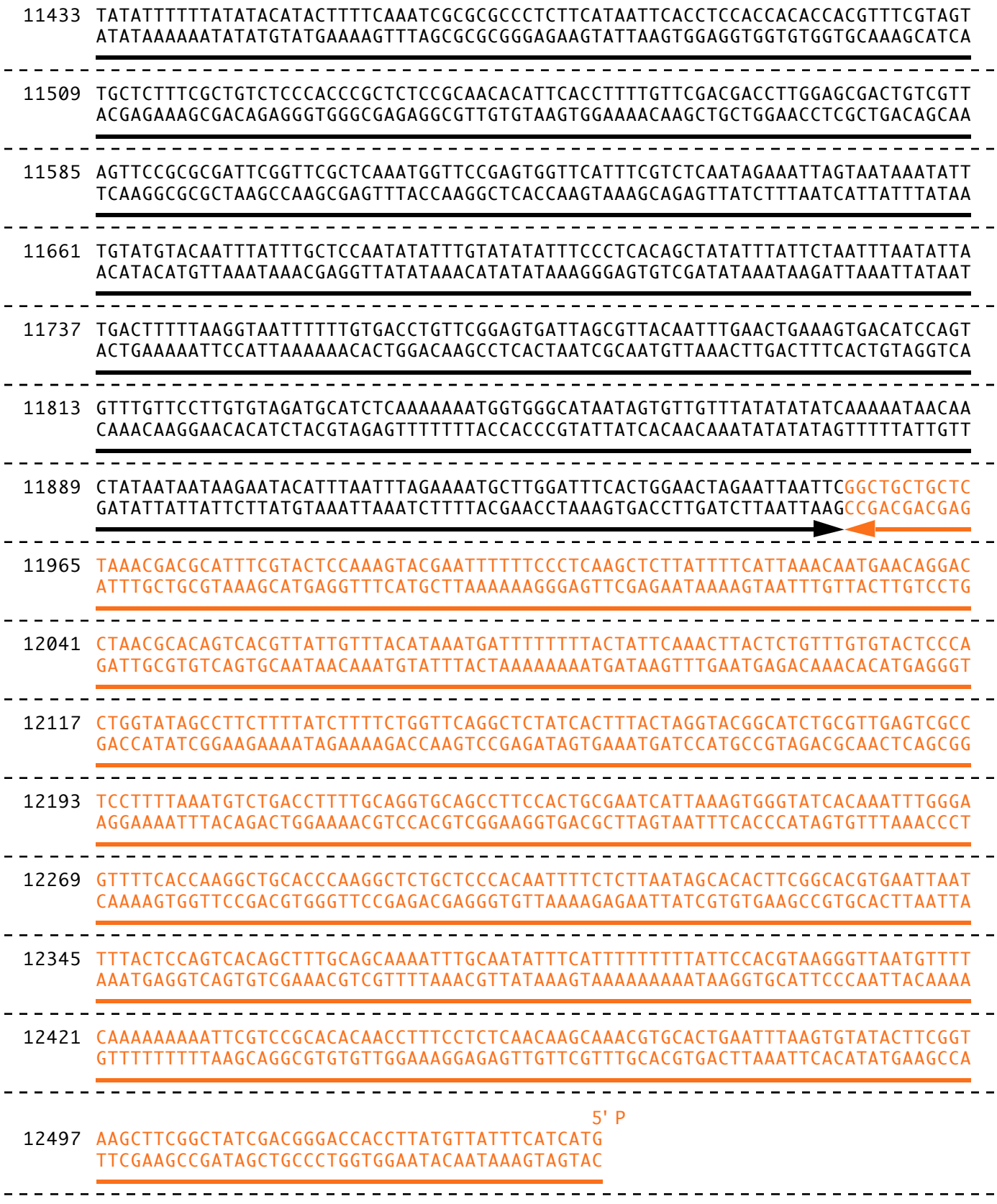

BglII
KpnI
NcoI
BamHI
attB2
attL2
EGFPN1 polylinker
KpnI
oligo 445 T7
EGFPN1
#54

```
  6372 TGG CCC ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC
       ACC GGG TGG GAG CAC TGG TGG GAC TGG ATG CCG CAC GTC ACG AAG TCG GCG ATG GGG

  6429 GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG
       CTG GTG TAC TTC GTC GTG CTG AAG AAG TTC AGG CGG TAC GGG CTT CCG ATG CAG GTC

  6486 GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG GTG AAG
       CTC GCG TGG TAG AAG AAG TTC CTG CTG CCG TTG ATG TTC TGG GCG CGG CTC CAC TTC

  6543 TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG
       AAG CTC CCG CTG TGG GAC CAC TTG GCG TAG CTC GAC TTC CCG TAG CTG AAG TTC CTC

  6600 GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT
       CTG CCG TTG TAG GAC CCC GTG TTC GAC CTC ATG TTG ATG TTG TCG GTG TTG CAG ATA

  6657 ATC ATG GCC GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC
       TAG TAC CGG CTG TTC GTC TTC TTG CCG TAG TTC CAC TTG AAG TTC TAG GCG GTG TTG

  6714 ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC
       TAG CTC CTG CCG TCG CAC GTC GAG CGG CTG GTG ATG GTC GTC TTG TGG GGG TAG CCG

  6771 GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC
       CTG CCG GGG CAC GAC GAC GGG CTG TTG GTG ATG GAC TCG TGG GTC AGG CGG GAC TCG

  6828 AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC
       TTT CTG GGG TTG CTC TTC GCG CTA GTG TAC CAG GAC GAC CTC AAG CAC TGG CGG CGG

                                                                    XbaI SV40 Poly A
  6885 GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TAA AGCGGCCGCGACTCTAGAGGATCTTTGT
       CCC TAG TGA GAG CCG TAC CTG CTC GAC ATG TTC ATT TCGCCGGCGCTGAGATCTCCTAGAAACA

  6949 GAAGGAACCTTACTTCTGTGGTGTGACATAATTGGACAAACTACCTACAGAGATTTAAAGCTCTAAGGTAAATATA
       CTTCCTTGGAATGAAGACACCACACTGTATTAACCTGTTTGATGGATGTCTCTAAATTTCGAGATTCCATTTATAT

  7025 AAATTTTTAAGTGTATAATGTGTTAAACTACTGATTCTAATTGTTTGTGTATTTTAGATTCCAACCTATGGAACTG
       TTTAAAAATTCACATATTACACAATTTGATGACTAAGATTAACAAACACATAAAATCTAAGGTTGGATACCTTGAC

  7101 ATGAATGGGAGCAGTGGTGGAATGCCTTTAATGAGGAAAACCTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGA
       TACTTACCCTCGTCACCACCTTACGGAAATTACTCCTTTTGGACAAAACGAGTCTTCTTTACGGTAGATCACTACT

  7177 TGAGGCTACTGCTGACTCTCAACATTCTACTCCTCCAAAAAAGAAGAGAAAGGTAGAAGACCCCAAGGACTTTCCT
       ACTCCGATGACGACTGAGAGTTGTAAGATGAGGAGGTTTTTTCTTCTCTTTCCATCTTCTGGGGTTCCTGAAAGGA

  7253 TCAGAATTGCTAAGTTTTTTGAGTCATGCTGTGTTTAGTAATAGAACTCTTGCTTGCTTTGCTATTTACACCACAA
       AGTCTTAACGATTCAAAAAACTCAGTACGACACAAATCATTATCTTGAGAACGAACGAAACGATAAATGTGGTGTT

  7329 AGGAAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATATTCTGTAACCTTTATAAGTAGGCATAACAGTTA
       TCCTTTTTCGACGTGACGATATGTTCTTTTAATACCTTTTTATAAGACATTGGAAATATTCATCCGTATTGTCAAT

  7405 TAATCATAACATACTGTTTTTTCTTACTCCACACAGGCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAATTG
       ATTAGTATTGTATGACAAAAAAGAATGAGGTGTGTCCGTATCTCACAGACGATAATTATTGATACGAGTTTTTAAC

  7481 TGTACCTTTAGCTTTTTAATTTGTAAAGGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAGATCATA
       ACATGGAAATCGAAAAATTAAACATTTCCCCAATTATTCCTTATAAACTACATATCACGGAACTGATCTCTAGTAT

  7557 ATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATA
       TAGTCGGTATGGTGTAAACATCTCCAAAATGAACGAAATTTTTTGGAGGGTGTGGAGGGGGACTTGGACTTTGTAT
```


HpaI
7633
AAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAA
TTTACTTACGTTAACAACAACAATTGAACAAATAACGTCGAATATTACCAATGTTTATTTCGTTATCGTAGTGTTT
7709
TTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTC
AAAGTGTTTATTTCGTAAAAAAAGTGACGTAAGATCAACACCAAACAGGTTTGAGTAGTTACATAGAATAGTACAG
white gene
BamHI
7785
TGGATCGGATCCACTAGAAGGCCTTAGTATGTATGTAAGTTAATAAAACCCTTTTTTGGAGAATGTAGATTTAAAA
ACCTAGCCTAGGTGATCTTCCGGAATCATACATACATTCAATTATTTTGGGAAAAAACCTCTTACATCTAAATTTT
7861
AAACATATTTTTTTTTTATTTTTTACTGCACTGGACATCATTGAACTTATCTGATCAGTTTTAAATTTACTTCGAT
TTTGTATAAAAAAAAAATAAAAAATGACGTGACCTGTAGTAACTTGAATAGACTAGTCAAAATTTAAATGAAGCTA
7937
CCAAGGGTATTTGAAGTACCAGGTTCTTTCGATTACCTCTCACTCAAAATGACATTCCACTCAAAGTCAGCGCTGT
GGTTCCCATAAACTTCATGGTCCAAGAAAGCTAATGGAGAGTGAGTTTTACTGTAAGGTGAGTTTCAGTCGCGACA
8013
TTGCCTCCTTCTCTGTCCACAGAAATATCGCCGTCTCTTTCGCCGCTGCGTCCGCTATCTCTTTCGCCACCGTTTG
AACGGAGGAAGAGACAGGTGTCTTTATAGCGGCAGAGAAAGCGGCGACGCAGGCGATAGAGAAAGCGGTGGCAAAC
8089
TAGCGTTACCTAGCGTCAATGTCCGCCTTCAGTTGCACTTTGTCAGCGGTTTCGTGACGAAGCTCCAAGCGGTTTA
ATCGCAATGGATCGCAGTTACAGGCGGAAGTCAACGTGAAACAGTCGCCAAAGCACTGCTTCGAGGTTCGCCAAAT
8165
CGCCATCAATTAAACACAAAGTGCTGTGCCAAAACTCCTCTCGCTTCTTATTTTTGTTTGTTTTTTGAGTGATTGG
GCGGTAGTTAATTTGTGTTTCACGACACGGTTTTGAGGAGAGCGAAGAATAAAAACAAACAAAAAACTCACTAACC
8241
GGTGGTGATTGGTTTTGGGTGGGTAAGCAGGGGAAAGTGTGAAAAATCCCGGCAATGGGCCAAGAGGATCAGGAGC
CCACCACTAACCAAAACCCACCCATTCGTCCCCTTTCACACTTTTTAGGGCCGTTACCCGGTTCTCCTAGTCCTCG
8317
TATTAATTCGCGGAGGCAGCAAACACCCATCTGCCGAGCATCTGAACAATGTGAGTAGTACATGTGCATACATCTT
ATAATTAAGCGCCTCCGTCGTTTGTGGGTAGACGGCTCGTAGACTTGTTACACTCATCATGTACACGTATGTAGAA
8393
AAGTTCACTTGATCTATAGGAACTGCGATTGCAACATCAAATTGTCTGCGGCGTGAGAACTGCGACCCACAAAAAT
TTCAAGTGAACTAGATATCCTTGACGCTAACGTTGTAGTTTAACAGACGCCGCACTCTTGACGCTGGGTGTTTTTA
8469
CCCAAACCGCAATCGCACAAACAAATAGTGACACGAAACAGATTATTCTGGTAGCTGTGCTCGCTATATAAGACAA
GGGTTTGGCGTTAGCGTGTTTGTTTATCACTGTGCTTTGTCTAATAAGACCATCGACACGAGCGATATATTCTGTT
8545
TTTTTAAGATCATATCATGATCAAGACATCTAAAGGCATTCATTTTCGACTACATTCTTTTTTACAAAAAATATAA
AAAAATTCTAGTATAGTACTAGTTCTGTAGATTTCCGTAAGTAAAAGCTGATGTAAGAAAAAATGTTTTTTATATT
8621
CAACCAGATATTTTAAGCTGATCCTAGATGCACAAAAAATAAATAAAAGTATAAACCTACTTCGTAGGATACTTCG
GTTGGTCTATAAAATTCGACTAGGATCTACGTGTTTTTTATTTATTTTCATATTTGGATGAAGCATCCTATGAAGC
8697
TTTTGTTCGGGGTTAGATGAGCATAACGCTTGTAGTTGATATTTGAGATCCCCTATCATTGCAGGGTGACAGCGGA
AAAACAAGCCCCAATCTACTCGTATTGCGAACATCAACTATAAACTCTAGGGGATAGTAACGTCCCACTGTCGCCT
8773
GCGGCTTCGCAGAGCTGCATTAACCAGGGCTTCGGGCAGGCCAAAAACTACGGCACGCTCCTGCCACCCAGTCCGC
CGCCGAAGCGTCTCGACGTAATTGGTCCCGAAGCCCGTCCGGTTTTTGATGCCGTGCGAGGACGGTGGGTCAGGCG

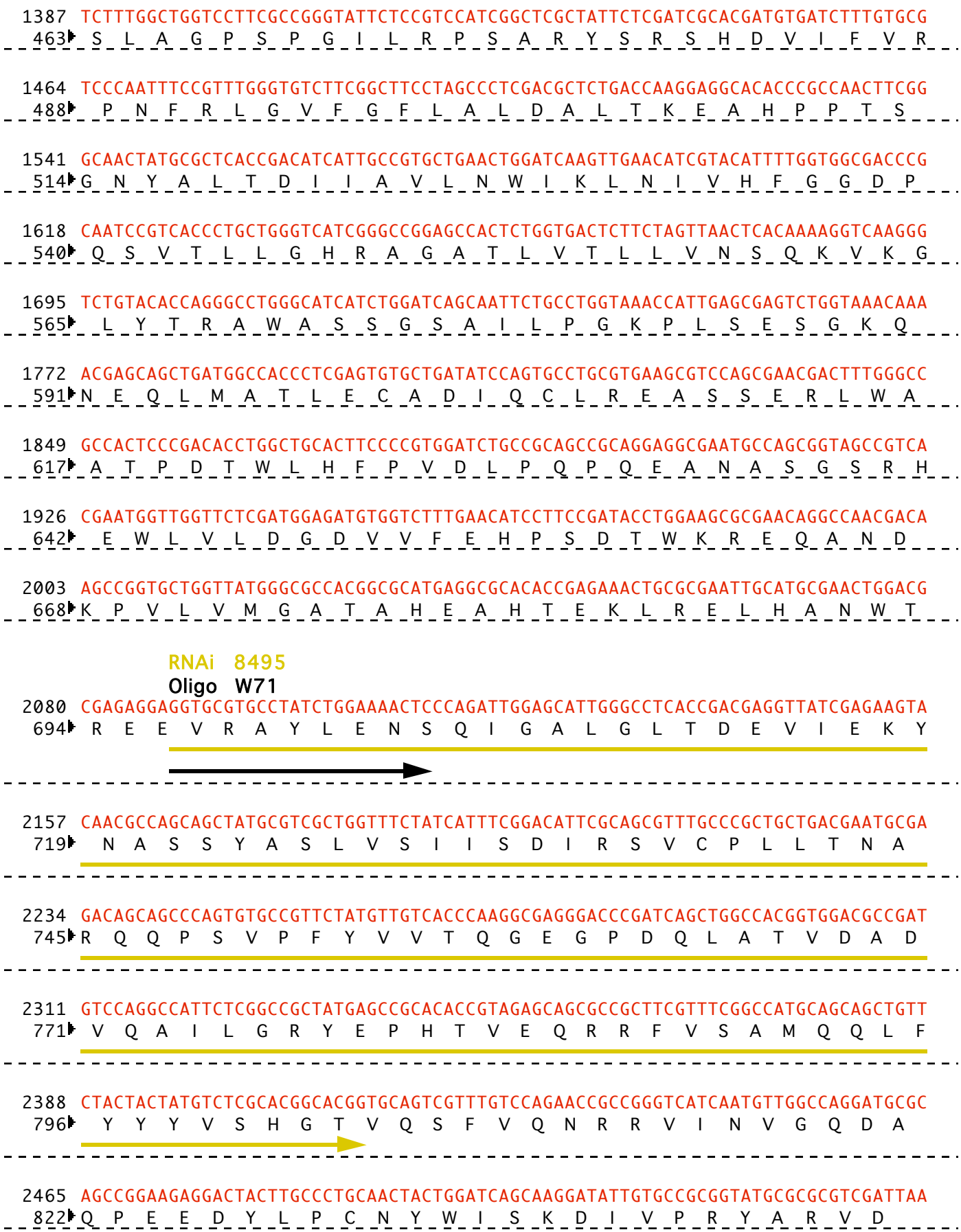
```
  8849 CGGAGGACTCCGGTTCAGGGAGCGGCCAACTAGCCGAGAACCTCACCTATGCCTGGCACAATATGGACATCTTTGG
       GCCTCCTGAGGCCAAGTCCCTCGCCGGTTGATCGGCTCTTGGAGTGGATACGGACCGTGTTATACCTGTAGAAACC

  8925 GGCGGTCAATCAGCCGGGCTCCGGATGGCGGCAGCTGGTCAACCGGACACGCGGACTATTCTGCAACGAGCGACAC
       CCGCCAGTTAGTCGGCCCGAGGCCTACCGCCGTCGACCAGTTGGCCTGTGCGCCTGATAAGACGTTGCTCGCTGTG

  9001 ATACCGGCGCCCAGGAAACATTTGCTCAAGAACGGTGAGTTTCTATTCGCAGTCGGCTGATCTGTGTGAAATCTTA
       TATGGCCGCGGGTCCTTTGTAAACGAGTTCTTGCCACTCAAAGATAAGCGTCAGCCGACTAGACACACTTTAGAAT

  9077 ATAAAGGGTCCAATTACCAATTTGAAACTCAGTTTGCGGCGTGGCCTATCCGGGCGAACTTTTGGCCGTGATGGGC
       TATTTCCCAGGTTAATGGTTAAACTTTGAGTCAAACGCCGCACCGGATAGGCCCGCTTGAAAACCGGCACTACCCG

  9153 AGTTCCGGTGCCGGAAAGACGACCCTGCTGAATGCCCTTGCCTTTCGATCGCCGCAGGGCATCCAAGTATCGCCAT
       TCAAGGCCACGGCCTTTCTGCTGGGACGACTTACGGGAACGGAAAGCTAGCGGCGTCCCGTAGGTTCATAGCGGTA

  9229 CCGGGATGCGACTGCTCAATGGCCAACCTGTGGACGCCAAGGAGATGCAGGCCAGGTGCGCCTATGTCCAGCAGGA
       GGCCCTACGCTGACGAGTTACCGGTTGGACACCTGCGGTTCCTCTACGTCCGGTCCACGCGGATACAGGTCGTCCT

  9305 TGACCTCTTTATCGGCTCCCTAACGGCCAGGGAACACCTGATTTTCCAGGCCATGGTGCGGATGCCACGACATCTG
       ACTGGAGAAATAGCCGAGGGATTGCCGGTCCCTTGTGGACTAAAAGGTCCGGTACCACGCCTACGGTGCTGTAGAC

  9381 ACCTATCGGCAGCGAGTGGCCCGCGTGGATCAGGTGATCCAGGAGCTTTCGCTCAGCAAATGTCAGCACACGATCA
       TGGATAGCCGTCGCTCACCGGGCGCACCTAGTCCACTAGGTCCTCGAAAGCGAGTCGTTTACAGTCGTGTGCTAGT

  9457 TCGGTGTGCCCGGCAGGGTGAAAGGTCTGTCCGGCGGAGAAAGGAAGCGTCTGGCATTCGCCTCCGAGGCACTAAC
       AGCCACACGGGCCGTCCCACTTTCCAGACAGGCCGCCTCTTTCCTTCGCAGACCGTAAGCGGAGGCTCCGTGATTG

  9533 CGATCCGCCGCTTCTGATCTGCGATGAGCCCACCTCCGGACTGGACTCATTTACCGCCCACAGCGTCGTCCAGGTG
       GCTAGGCGGCGAAGACTAGACGCTACTCGGGTGGAGGCCTGACCTGAGTAAATGGCGGGTGTCGCAGCAGGTCCAC

  9609 CTGAAGAAGCTGTCGCAGAAGGGCAAGACCGTCATCCTGACCATTCATCAGCCGTCTTCCGAGCTGTTTGAGCTCT
       GACTTCTTCGACAGCGTCTTCCCGTTCTGGCAGTAGGACTGGTAAGTAGTCGGCAGAAGGCTCGACAAACTCGAGA

  9685 TTGACAAGATCCTTCTGATGGCCGAGGGCAGGGTAGCTTTCTTGGGCACTCCCAGCGAAGCCGTCGACTTCTTTTC
       AACTGTTCTAGGAAGACTACCGGCTCCCGTCCCATCGAAAGAACCCGTGAGGGTCGCTTCGGCAGCTGAAGAAAAG

  9761 CTAGTGAGTTCGATGTGTTTATTAAGGGTATCTAGCATTACATTACATCTCAACTCCTATCCAGCGTGGGTGCCCA
       GATCACTCAAGCTACACAAATAATTCCCATAGATCGTAATGTAATGTAGAGTTGAGGATAGGTCGCACCCACGGGT

  9837 GTGTCCTACCAACTACAATCCGGCGGACTTTTACGTACAGGTGTTGGCCGTTGTGCCCGGACGGGAGATCGAGTCC
       CACAGGATGGTTGATGTTAGGCCGCCTGAAAATGCATGTCCACAACCGGCAACACGGGCCTGCCCTCTAGCTCAGG

  9913 CGTGATCGGATCGCCAAGATATGCGACAATTTTGCTATTAGCAAAGTAGCCCGGGATATGGAGCAGTTGTTGGCCA
       GCACTAGCCTAGCGGTTCTATACGCTGTTAAAACGATAATCGTTTCATCGGGCCCTATACCTCGTCAACAACCGGT

  9989 CCAAAAATTTGGAGAAGCCACTGGAGCAGCCGGAGAATGGGTACACCTACAAGGCCACCTGGTTCATGCAGTTCCG
       GGTTTTTAAACCTCTTCGGTGACCTCGTCGGCCTCTTACCCATGTGGATGTTCCGGTGGACCAAGTACGTCAAGGC

 10065 GGCGGTCCTGTGGCGATCCTGGCTGTCGGTGCTCAAGGAACCACTCCTCGTAAAAGTGCGACTTATTCAGACAACG
       CCGCCAGGACACCGCTAGGACCGACAGCCACGAGTTCCTTGGTGAGGAGCATTTTCACGCTGAATAAGTCTGTTGC
```


10141 GTGAGTGGTTCCAGTGGAAACAAATGATATAACGCTTACAATTCTTGGAAACAAATTCGCTAGATTTTAGTTAGAA
CACTCACCAAGGTCACCTTTGTTTACTATATTGCGAATGTTAAGAACCTTTGTTTAAGCGATCTAAAATCAATCTT
10217 TTGCCTGATTCCACACCCTTCTTAGTTTTTTTCAATGAGATGTATAGTTTATAGTTTTGCAGAAAATAAATAAATT
AACGGACTAAGGTGTGGGAAGAATCAAAAAAAGTTACTCTACATATCAAATATCAAAACGTCTTTTATTTATTTAA
10293 TCATTTAACTCGCGAACATGTTGAAGATATGAATATTAATGAGATGCGAGTAACATTTTAATTTGCAGATGGTTGC
AGTAAATTGAGCGCTTGTACAACTTCTATACTTATAATTACTCTACGCTCATTGTAAAATTAAACGTCTACCAACG
10369 CATCTTGATTGGCCTCATCTTTTTGGGCCAACAACTCACGCAAGTGGGCGTGATGAATATCAACGGAGCCATCTTC
GTAGAACTAACCGGAGTAGAAAAACCCGGTTGTTGAGTGCGTTCACCCGCACTACTTATAGTTGCCTCGGTAGAAG
10445 CTCTTCCTGACCAACATGACCTTTCAAAACGTCTTTGCCACGATAAATGTAAGTCTTGTTTAGAATACATTTGCAT
GAGAAGGACTGGTTGTACTGGAAAGTTTTGCAGAAACGGTGCTATTTACATTCAGAACAAATCTTATGTAAACGTA
10521 ATTAATAATTTACTAACTTTCTAATGAATCGATTCGATTTAGGTGTTCACCTCAGAGCTGCCAGTTTTTATGAGGG
TAATTATTAAATGATTGAAAGATTACTTAGCTAAGCTAAATCCACAAGTGGAGTCTCGACGGTCAAAAATACTCCC
10597 AGGCCCGAAGTCGACTTTATCGCTGTGACACATACTTTCTGGGCAAAACGATTGCCGAATTACCGCTTTTTCTCAC
TCCGGGCTTCAGCTGAAATAGCGACACTGTGTATGAAAGACCCGTTTTGCTAACGGCTTAATGGCGAAAAAGAGTG
10673 AGTGCCACTGGTCTTCACGGCGATTGCCTATCCGATGATCGGACTGCGGGCCGGAGTGCTGCACTTCTTCAACTGC
TCACGGTGACCAGAAGTGCCGCTAACGGATAGGCTACTAGCCTGACGCCCGGCCTCACGACGTGAAGAAGTTGACG
10749 CTGGCGCTGGTCACTCTGGTGGCCAATGTGTCAACGTCCTTCGGATATCTAATATCCTGCGCCAGCTCCTCGACCT
GACCGCGACCAGTGAGACCACCGGTTACACAGTTGCAGGAAGCCTATAGATTATAGGACGCGGTCGAGGAGCTGGA
10825 CGATGGCGCTGTCTGTGGGTCCGCCGGTTATCATACCATTCCTGCTCTTTGGCGGCTTCTTCTTGAACTCGGGCTC
GCTACCGCGACAGACACCCAGGCGGCCAATAGTATGGTAAGGACGAGAAACCGCCGAAGAAGAACTTGAGCCCGAG
10901 GGTGCCAGTATACCTCAAATGGTTGTCGTACCTCTCATGGTTCCGTTACGCCAACGAGGGTCTGCTGATTAACCAA
CCACGGTCATATGGAGTTTACCAACAGCATGGAGAGTACCAAGGCAATGCGGTTGCTCCCAGACGACTAATTGGTT
10977 TGGGCGGACGTGGAGCCGGGCGAAATTAGCTGCACATCGTCGAACACCACGTGCCCCAGTTCGGGCAAGGTCATCC
ACCCGCCTGCACCTCGGCCCGCTTTAATCGACGTGTAGCAGCTTGTGGTGCACGGGGTCAAGCCCGTTCCAGTAGG
11053 TGGAGACGCTTAACTTCTCCGCCGCCGATCTGCCGCTGGACTACGTGGGTCTGGCCATTCTCATCGTGAGCTTCCG
ACCTCTGCGAATTGAAGAGGCGGCGGCTAGACGGCGACCTGATGCACCCAGACCGGTAAGAGTAGCACTCGAAGGC
11129 GGTGCTCGCATATCTGGCTCTAAGACTTCGGGCCCGACGCAAGGAGTAGCCGACATATATCCGAAATAACTGCTTG
CCACGAGCGTATAGACCGAGATTCTGAAGCCCGGGCTGCGTTCCTCATCGGCTGTATATAGGCTTTATTGACGAAC
11205 TTTTTTTTTTTACCATTATTACCATCGTGTTTACTGTTTATTGCCCCCTCAAAAAGCTAATGTAATTATATTTGTG
AAAAAAAAAAAATGGTAATAATGGTAGCACAAATGACAAATAACGGGGGAGTTTTTCGATTACATTAATATAAACAC
11281 CCAATAAAAACAAGATATGACCTATAGAATACAAGTATTTCCCCTTCGAACATCCCCACAAGTAGACTTTGGATTT
GGTTATTTTTGTTCTATACTGGATATCTTATGTTCATAAAGGGGAAGCTTGTAGGGGTGTTCATCTGAAACCTAAA
11357 GTCTTCTAACCAAAAGACTTACACACCTGCATACCTTACATCAAAAACTCGTTTATCGCTACATAAAACACCGGGA
CAGAAGATTGGTTTTCTGAATGTGTGGACGTATGGAATGTAGTTTTTGAGCAAATAGCGATGTATTTTGTGGCCCT

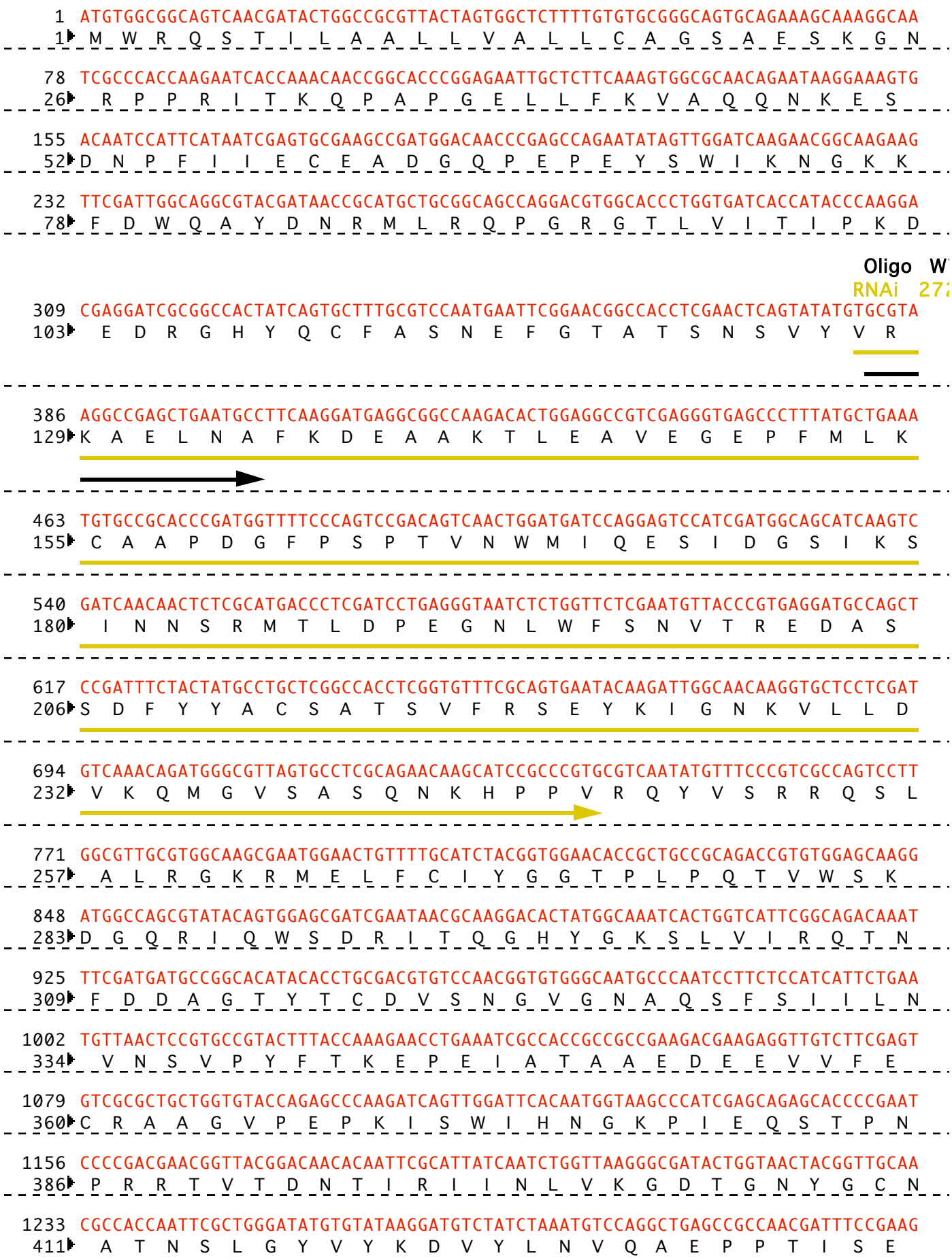

11433 TATATTTTTTATATACATACTTTTCAAATCGCGCGCCCTCTTCATAATTCACCTCCACCACACCACGTTTCGTAGT
ATATAAAAAATATATGTATGAAAAGTTTAGCGCGCGGGAGAAGTATTAAGTGGAGGTGGTGTGGTGCAAAGCATCA
11509 TGCTCTTTCGCTGTCTCCCACCCGCTCTCCGCAACACATTCACCTTTTGTTCGACGACCTTGGAGCGACTGTCGTT
ACGAGAAAGCGACAGAGGGTGGGCGAGAGGCGTTGTGTAAGTGGAAAACAAGCTGCTGGAACCTCGCTGACAGCAA
11585 AGTTCCGCGCGATTCGGTTCGCTCAAATGGTTCCGAGTGGTTCATTTCGTCTCAATAGAAATTAGTAATAAATATT
TCAAGGCGCGCTAAGCCAAGCGAGTTTACCAAGGCTCACCAAGTAAAGCAGAGTTATCTTTAATCATTATTTATAA
11661 TGTATGTACAATTTATTTGCTCCAATATATTTGTATATATTTCCCTCACAGCTATATTTATTCTAATTTAATATTA
ACATACATGTTAAATAAACGAGGTTATATAAACATATATAAAGGGAGTGTCGATATAAATAAGATTAAATTATAAT
11737 TGACTTTTTAAGGTAATTTTTTGTGACCTGTTCGGAGTGATTAGCGTTACAATTTGAACTGAAAGTGACATCCAGT
ACTGAAAAATTCCATTAAAAAACACTGGACAAGCCTCACTAATCGCAATGTTAAACTTGACTTTCACTGTAGGTCA
11813 GTTTGTTCCTTGTGTAGATGCATCTCAAAAAAATGGTGGGCATAATAGTGTTGTTTATATATATCAAAAATAACAA
CAAACAAGGAACACATCTACGTAGAGTTTTTTTACCACCCGTATTATCACAACAAATATATATAGTTTTTATTGTT
11889 CTATAATAATAAGAATACATTTAATTTAGAAAATGCTTGGATTTCACTGGAACTAGAATTAATTCGGCTGCTGCTC
GATATTATTATTCTTATGTAAATTAAATCTTTTACGAACCTAAAGTGACCTTGATCTTAATTAAGCCGACGACGAG
11965 TAAACGACGCATTTCGTACTCCAAAGTACGAATTTTTTCCCTCAAGCTCTTATTTTCATTAAACAATGAACAGGAC
ATTTGCTGCGTAAAGCATGAGGTTTCATGCTTAAAAAAGGGAGTTCGAGAATAAAAGTAATTTGTTACTTGTCCTG
12041 CTAACGCACAGTCACGTTATTGTTTACATAAATGATTTTTTTTACTATTCAAACTTACTCTGTTTGTGTACTCCCA
GATTGCGTGTCAGTGCAATAACAAATGTATTTACTAAAAAAAAATGATAAGTTTGAATGAGACAAACACATGAGGGT
12117 CTGGTATAGCCTTCTTTTATCTTTTCTGGTTCAGGCTCTATCACTTTACTAGGTACGGCATCTGCGTTGAGTCGCC
GACCATATCGGAAGAAAATAGAAAAGACCAAGTCCGAGATAGTGAAATGATCCATGCCGTAGACGCAACTCAGCGG
12193 TCCTTTTAAATGTCTGACCTTTTGCAGGTGCAGCCTTCCACTGCGAATCATTAAAGTGGGTATCACAAATTTGGGA
AGGAAAATTTACAGACTGGAAAACGTCCACGTCGGAAGGTGACGCTTAGTAATTTCACCCATAGTGTTTAAACCCT
12269 GTTTTCACCAAGGCTGCACCCAAGGCTCTGCTCCCACAATTTTCTCTTAATAGCACACTTCGGCACGTGAATTAAT
CAAAAGTGGTTCCGACGTGGGTTCCGAGACGAGGGTGTTAAAAGAGAATTATCGTGTGAAGCCGTGCACTTAATTA
12345 TTTACTCCAGTCACAGCTTTGCAGCAAAATTTGCAATATTTCATTTTTTTTTATTCCACGTAAGGGTTAATGTTTT
AAATGAGGTCAGTGTCGAAACGTCGTTTTAAACGTTATAAAGTAAAAAAAAAATAAGGTGCATTCCCAATTACAAAA
12421 CAAAAAAAAATTCGTCCGCACACAACCTTTCCTCTCAACAAGCAAACGTGCACTGAATTTAAGTGTATACTTCGGT
GTTTTTTTTTAAGCAGGCGTGTGTTGGAAAGGAGAGTTGTTCGTTTGCACGTGACTTAAATTCACATATGAAGCCA
5' P
12497 AAGCTTCGGCTATCGACGGGACCACCTTATGTTATTTCATCATG
TTCGAAGCCGATAGCTGCCCTGGTGGAATACAATAAAGTAGTAC

NRT ORF

```
    1 ATGGGCGAACTCGAGGAGAAGGAAACCCCGCCCACTGAGACGACAGCCGCCCAGCAAGAGGCGCTAGAGGAGCCCAA
    1 M  G  E  L  E  E  K  E  T  P  P  T  E  T  T  A  A  Q  Q  E  A  L  E  E  P  K

   78 GGAAACGGACAAAATGTTGGACAAAAAAGAGGACGCCAAGGAGAAGACACCCAGTCCACAGACCTCCAAGCCCGCAT
   26  E  T  D  K  M  L  D  K  K  E  D  A  K  E  K  T  P  S  P  Q  T  S  K  P  A

  155 CTCCAAATGCCGGCAAGAAATCCTCACCAGTGGCCGAGAAAAAGATCGACGATGCTGAATTAGCGAAATCCAAATCA
   52 S  P  N  A  G  K  K  S  S  P  V  A  E  K  K  I  D  D  A  E  L  A  K  S  K  S

  232 GGCAATGGAGAAGAGATTATCGATATTCCCGCCGAGAATGGCACAAAGCCAGACAGCGCTGATGACAAAAAGATAAG
   78 G  N  G  E  E  I  I  D  I  P  A  E  N  G  T  K  P  D  S  A  D  D  K  K  I  S

  309 CAAGGAGGAGCGCGAGGTCAAGCCCAAAAAGATACCGATCGGAGGTCTCAAACTGCCTGGTTTCTTCATGAAGAACA
  103  K  E  E  R  E  V  K  P  K  K  I  P  I  G  G  L  K  L  P  G  F  F  M  K  N

  386 AGCCGAAGGCAGATGGTGATGGGGCCGAGGGCGAGCTGCTCGAAAAGGAGAAGGAAGAGGATAAGGATAAGGAAGCC
  129 K  P  K  A  D  G  D  G  A  E  G  E  L  L  E  K  E  K  E  E  D  K  D  K  E  A

  463 AATGGAGATGCCGCCACCGGTTCCGGCAAGGACGAACAGAAATCTCGCCCAGGACTGGGAGAACGCCTGCGCAGCTT
  155 N  G  D  A  A  T  G  S  G  K  D  E  Q  K  S  R  P  G  L  G  E  R  L  R  S  F

  540 CTTTGCCCGCAAGCCATCCGCCGAAAAGGAAAAGAAGCAGCTGGTCAACGGTGACGCGGATGCCAAGTCTGAAGCCA
  180  F  A  R  K  P  S  A  E  K  E  K  K  Q  L  V  N  G  D  A  D  A  K  S  E  A

  617 CAGCTGAAGCAACGCCCGCTGAAGATGCCTCCGATGCACCACCAAAGCGTGGACTTTTGAACGCCATCAAGCTGCCA
  206 T  A  E  A  T  P  A  E  D  A  S  D  A  P  P  K  R  G  L  L  N  A  I  K  L  P

  694 ATCGCTAACATGATACCGAAAAAGAAGAGCAACGATGATGTGGAGCTGGGCTTGGGCAAGGCCGGTCTGGCCTCGAT
  232 I  A  N  M  I  P  K  K  K  S  N  D  D  V  E  L  G  L  G  K  A  G  L  A  S  M

  771 GGAGACCCTCGATGATTCCCTTAAGGATCAAGACACAGTGGATCGGGCTCCCGTCAAGACCAACGGTACCGAGGAAC
  257  E  T  L  D  D  S  L  K  D  Q  D  T  V  D  R  A  P  V  K  T  N  G  T  E  E

  848 TAAAGGGCGAGCTAAAGGATGAGAAGCTGGCGGCGGAGGAAAAACTAGCCGCCGAGGAGGAGGAGCAAAACCGACCC
  283 L  K  G  E  L  K  D  E  K  L  A  A  E  E  K  L  A  A  E  E  E  E  Q  N  R  P

  925 GTCTCCTTGCTAACCCGTCTGCGTGGCTACAAGTGCAGTGTGGACGATGCCCTGATTGTGTTTGGCATCCTGCTATT
  309 V  S  L  L  T  R  L  R  G  Y  K  C  S  V  D  D  A  L  I  V  F  G  I  L  L  F

 1002 TGTGCTCCTGTTGGGCGTGATTGGTTATGTACTAACCCACGAGACTTTGACCTCGCCGCCGCTGCGGGAAGGACGCT
  334  V  L  L  L  G  V  I  G  Y  V  L  T  H  E  T  L  T  S  P  P  L  R  E  G  R

 1079 ACATAATGGCAGTGACGGGGTGCGGACCTGTGGAGGGCGTTAAGGAAGATGGAGCCTTTGCCTTCCGTGGCATTCCG
  360 Y  I  M  A  V  T  G  C  G  P  V  E  G  V  K  E  D  G  A  F  A  F  R  G  I  P

 1156 TATGCAAAGCCACCCGTAGACAGACTGAGATGGAAGCCGGCTGAACTGATTGATGACATCAATATGTGCTGGAATGA
  386 Y  A  K  P  P  V  D  R  L  R  W  K  P  A  E  L  I  D  D  I  N  M  C  W  N  D

 1233 TACACTGCAAACCCATAACAGCAGTGTGGTGTGCACGCAGCGATTGGGCAATGGCACCACAGTTGGCGACGAGGATT
  411  T  L  Q  T  H  N  S  S  V  V  C  T  Q  R  L  G  N  G  T  T  V  G  D  E  D

 1310 GTCTATACCTTGACGTGGTTACTCCCCATGTGCGGTACAATAACCCCTTGCCTGTGGTCGTCCTGATCGGAGCAGAA
  437 C  L  Y  L  D  V  V  T  P  H  V  R  Y  N  N  P  L  P  V  V  V  L  I  G  A  E
```


1387 TCTTTGGCTGGTCCTTCGCCGGGTATTCTCCGTCCATCGGCTCGCTATTCTCGATCGCACGATGTGATCTTTGTGCG
463 S L A G P S P G I L R P S A R Y S R S H D V I F V R
1464 TCCCAATTTCCGTTTGGGTGTCTTCGGCTTCCTAGCCCTCGACGCTCTGACCAAGGAGGCACACCCGCCAACTTCGG
488 P N F R L G V F G F L A L D A L T K E A H P P T S
1541 GCAACTATGCGCTCACCGACATCATTGCCGTGCTGAACTGGATCAAGTTGAACATCGTACATTTTGGTGGCGACCCG
514 G N Y A L T D I I A V L N W I K L N I V H F G G D P
1618 CAATCCGTCACCCTGCTGGGTCATCGGGCCGGAGCCACTCTGGTGACTCTTCTAGTTAACTCACAAAAGGTCAAGGG
540 Q S V T L L G H R A G A T L V T L L V N S Q K V K G
1695 TCTGTACACCAGGGCCTGGGCATCATCTGGATCAGCAATTCTGCCTGGTAAACCATTGAGCGAGTCTGGTAAACAAA
565 L Y T R A W A S S G S A I L P G K P L S E S G K Q
1772 ACGAGCAGCTGATGGCCACCCTCGAGTGTGCTGATATCCAGTGCCTGCGTGAAGCGTCCAGCGAACGACTTTGGGCC
591 N E Q L M A T L E C A D I Q C L R E A S S E R L W A
1849 GCCACTCCCGACACCTGGCTGCACTTCCCCGTGGATCTGCCGCAGCCGCAGGAGGCGAATGCCAGCGGTAGCCGTCA
617 A T P D T W L H F P V D L P Q P Q E A N A S G S R H
1926 CGAATGGTTGGTTCTCGATGGAGATGTGGTCTTTGAACATCCTTCCGATACCTGGAAGCGCGAACAGGCCAACGACA
642 E W L V L D G D V V F E H P S D T W K R E Q A N D
2003 AGCCGGTGCTGGTTATGGGCGCCACGGCGCATGAGGCGCACACCGAGAAACTGCGCGAATTGCATGCGAACTGGACG
668 K P V L V M G A T A H E A H T E K L R E L H A N W T
RNAi 8495
Oligo W71
2080 CGAGAGGAGGTGCGTGCCTATCTGGAAAACTCCCAGATTGGAGCATTGGGCCTCACCGACGAGGTTATCGAGAAGTA
694 R E E V R A Y L E N S Q I G A L G L T D E V I E K Y
2157 CAACGCCAGCAGCTATGCGTCGCTGGTTTCTATCATTTCGGACATTCGCAGCGTTTGCCCGCTGCTGACGAATGCGA
719 N A S S Y A S L V S I I S D I R S V C P L L T N A
2234 GACAGCAGCCCAGTGTGCCGTTCTATGTTGTCACCCAAGGCGAGGGACCCGATCAGCTGGCCACGGTGGACGCCGAT
745 R Q Q P S V P F Y V V T Q G E G P D Q L A T V D A D
2311 GTCCAGGCCATTCTCGGCCGCTATGAGCCGCACACCGTAGAGCAGCGCCGCTTCGTTTCGGCCATGCAGCAGCTGTT
771 V Q A I L G R Y E P H T V E Q R R F V S A M Q Q L F
2388 CTACTACTATGTCTCGCACGGCACGGTGCAGTCGTTTGTCCAGAACCGCCGGGTCATCAATGTTGGCCAGGATGCGC
796 Y Y Y V S H G T V Q S F V Q N R R V I N V G Q D A
2465 AGCCGGAAGAGGACTACTTGCCCTGCAACTACTGGATCAGCAAGGATATTGTGCCGCGGTATGCGCGCGTCGATTAA
822 Q P E E D Y L P C N Y W I S K D I V P R Y A R V D

NRG ORF


1 ATGTGGCGGCAGTCAACGATACTGGCCGCGTTACTAGTGGCTCTTTTGTGTGCGGGCAGTGCAGAAAGCAAAGGCAA
1 M W R Q S T I L A A L L V A L L C A G S A E S K G N
78 TCGCCCACCAAGAATCACCAAACAACCGGCACCCGGAGAATTGCTCTTCAAAGTGGCGCAACAGAATAAGGAAAGTG
26 R P P R I T K Q P A P G E L L F K V A Q Q N K E S
155 ACAATCCATTCATAATCGAGTGCGAAGCCGATGGACAACCCGAGCCAGAATATAGTTGGATCAAGAACGGCAAGAAG
52 D N P F I I E C E A D G Q P E P E Y S W I K N G K K
232 TTCGATTGGCAGGCGTACGATAACCGCATGCTGCGGCAGCCAGGACGTGGCACCCTGGTGATCACCATACCCAAGGA
78 F D W Q A Y D N R M L R Q P G R G T L V I T I P K D
Oligo W
RNAi 27
309 CGAGGATCGCGGCCACTATCAGTGCTTTGCGTCCAATGAATTCGGAACGGCCACCTCGAACTCAGTATATGTGCGTA
103 E D R G H Y Q C F A S N E F G T A T S N S V Y V R
386 AGGCCGAGCTGAATGCCTTCAAGGATGAGGCGGCCAAGACACTGGAGGCCGTCGAGGGTGAGCCCTTTATGCTGAAA
129 K A E L N A F K D E A A K T L E A V E G E P F M L K
463 TGTGCCGCACCCGATGGTTTTCCCAGTCCGACAGTCAACTGGATGATCCAGGAGTCCATCGATGGCAGCATCAAGTC
155 C A A P D G F P S P T V N W M I Q E S I D G S I K S
540 GATCAACAACTCTCGCATGACCCTCGATCCTGAGGGTAATCTCTGGTTCTCGAATGTTACCCGTGAGGATGCCAGCT
180 I N N S R M T L D P E G N L W F S N V T R E D A S
617 CCGATTTCTACTATGCCTGCTCGGCCACCTCGGTGTTTCGCAGTGAATACAAGATTGGCAACAAGGTGCTCCTCGAT
206 S D F Y Y A C S A T S V F R S E Y K I G N K V L L D
694 GTCAAACAGATGGGCGTTAGTGCCTCGCAGAACAAGCATCCGCCCGTGCGTCAATATGTTTCCCGTCGCCAGTCCTT
232 V K Q M G V S A S Q N K H P P V R Q Y V S R R Q S L
771 GGCGTTGCGTGGCAAGCGAATGGAACTGTTTTGCATCTACGGTGGAACACCGCTGCCGCAGACCGTGTGGAGCAAGG
257 A L R G K R M E L F C I Y G G T P L P Q T V W S K
848 ATGGCCAGCGTATACAGTGGAGCGATCGAATAACGCAAGGACACTATGGCAAATCACTGGTCATTCGGCAGACAAAT
283 D G Q R I Q W S D R I T Q G H Y G K S L V I R Q T N
925 TTCGATGATGCCGGCACATACACCTGCGACGTGTCCAACGGTGTGGGCAATGCCCAATCCTTCTCCATCATTCTGAA
309 F D D A G T Y T C D V S N G V G N A Q S F S I I L N
1002 TGTTAACTCCGTGCCGTACTTTACCAAAGAACCTGAAATCGCCACCGCCGCCGAAGACGAAGAGGTTGTCTTCGAGT
334 V N S V P Y F T K E P E I A T A A E D E E V V F E
1079 GTCGCGCTGCTGGTGTACCAGAGCCCAAGATCAGTTGGATTCACAATGGTAAGCCCATCGAGCAGAGCACCCCGAAT
360 C R A A G V P E P K I S W I H N G K P I E Q S T P N
1156 CCCCGACGAACGGTTACGGACAACACAATTCGCATTATCAATCTGGTTAAGGGCGATACTGGTAACTACGGTTGCAA
386 P R R T V T D N T I R I I N L V K G D T G N Y G C N
1233 CGCCACCAATTCGCTGGGATATGTGTATAAGGATGTCTATCTAAATGTCCAGGCTGAGCCGCCAACGATTTCCGAAG
411 A T N S L G Y V Y K D V Y L N V Q A E P P T I S E

```
1310 CTCCAGCAGCTGTATCCACTGTCGATGGAAGGAATGTGACCATTAAGTGCAGGGTTAACGGTTCCCCCAAGCCTCTG
 437▶A  P  A  A  V  S  T  V  D  G  R  N  V  T  I  K  C  R  V  N  G  S  P  K  P  L

1387 GTTAAATGGCTAAGGGCCAGCAACTGGCTGACCGGAGGTCGTTACAATGTCCAAGCTAACGGTGACCTGGAGATCCA
 463▶V  K  W  L  R  A  S  N  W  L  T  G  G  R  Y  N  V  Q  A  N  G  D  L  E  I  Q

1464 AGATGTGACATTCTCGGATGCCGGCAAATACACATGCTATGCGCAGAACAAGTTTGGTGAAATTCAAGCCGATGGTT
 488▶ D  V  T  F  S  D  A  G  K  Y  T  C  Y  A  Q  N  K  F  G  E  I  Q  A  D  G

1541 CGCTGGTGGTCAAGGAGCATACGAGAATTACCCAAGAGCCGCAAAACTACGAGGTGGCCGCCGGACAATCGGCCACG
 514▶S  L  V  V  K  E  H  T  R  I  T  Q  E  P  Q  N  Y  E  V  A  A  G  Q  S  A  T

1618 TTCCGCTGTAACGAGGCCCACGACGATACGCTGGAGATTGAGATCGATTGGTGGAAGGATGGCCAGTCCATTGACTT
 540▶F  R  C  N  E  A  H  D  D  T  L  E  I  E  I  D  W  W  K  D  G  Q  S  I  D  F

1695 TGAGGCCCAGCCGCGATTCGTGAAGACCAATGATAATTCCCTGACGATTGCCAAGACAATGGAGTTGGATTCTGGCG
 565▶ E  A  Q  P  R  F  V  K  T  N  D  N  S  L  T  I  A  K  T  M  E  L  D  S  G

1772 AATATACGTGCGTGGCCCGGACGCGTTTGGATGAGGCAACGGCCAGGGCGAATTTGATTGTCCAGGATGTGCCGAAT
 591▶E  Y  T  C  V  A  R  T  R  L  D  E  A  T  A  R  A  N  L  I  V  Q  D  V  P  N

1849 GCACCAAAACTGACCGGCATCACCTGCCAGGCCGACAAGGCCGAGATCCACTGGGAACAGCAGGGTGACAATCGTTC
 617▶A  P  K  L  T  G  I  T  C  Q  A  D  K  A  E  I  H  W  E  Q  Q  G  D  N  R  S

1926 GCCCATTCTGCACTACACCATTCAGTTCAATACATCGTTCACGCCCGCCTCCTGGGATGCCGCCTACGAGAAGGTGC
 642▶ P  I  L  H  Y  T  I  Q  F  N  T  S  F  T  P  A  S  W  D  A  A  Y  E  K  V

2003 CCAACACGGACTCCTCGTTCGTCGTCCAGATGTCACCGTGGGCCAACTATACGTTCCGTGTGATTGCCTTCAACAAG
 668▶P  N  T  D  S  S  F  V  V  Q  M  S  P  W  A  N  Y  T  F  R  V  I  A  F  N  K

2080 ATCGGAGCCTCGCCGCCGTCGGCGCACAGCGATAGCTGCACCACCCAGCCGGATGTGCCCTTCAAGAATCCCGACAA
 694▶I  G  A  S  P  P  S  A  H  S  D  S  C  T  T  Q  P  D  V  P  F  K  N  P  D  N

2157 TGTCGTTGGCCAGGGCACTGAGCCCAACAATCTGGTCATCTCGTGGACTCCCATGCCCGAAATCGAGCACAATGCCC
 719▶ V  V  G  Q  G  T  E  P  N  N  L  V  I  S  W  T  P  M  P  E  I  E  H  N  A

2234 CCAATTTCCATTATTATGTTAGCTGGAAACGCGATATTCCTGCCGCTGCGTGGGAAAACAATAACATATTCGACTGG
 745▶P  N  F  H  Y  Y  V  S  W  K  R  D  I  P  A  A  A  W  E  N  N  N  I  F  D  W

2311 CGACAGAACAACATTGTGATTGCCGATCAACCGACTTTCGTGAAATACCTGATCAAGGTGGTGGCCATCAACGATAG
 771▶R  Q  N  N  I  V  I  A  D  Q  P  T  F  V  K  Y  L  I  K  V  V  A  I  N  D  R

2388 GGGTGAGTCCAATGTGGCCGCCGAGGAGGTGGTTGGCTACTCTGGCGAAGATCGTCCCCTGGATGCGCCCACCAACT
 796▶ G  E  S  N  V  A  A  E  E  V  V  G  Y  S  G  E  D  R  P  L  D  A  P  T  N

2465 TCACAATGAGGCAAATCACATCATCGACCAGTGGCTACATGGCCTGGACGCCGGTAAGTGAGGAATCGGTGCGCGGA
 822▶F  T  M  R  Q  I  T  S  S  T  S  G  Y  M  A  W  T  P  V  S  E  E  S  V  R  G

2542 CACTTCAAGGGCTACAAAATCCAAACGTGGACGGAGAACGAGGGCGAGGAGGGTCTGCGGGAGATCCATGTGAAGGG
 848▶H  F  K  G  Y  K  I  Q  T  W  T  E  N  E  G  E  E  G  L  R  E  I  H  V  K  G

2619 TGATACCCACAACGCTCTGGTCACACAATTCAAGCCCGATTCAAAGAACTATGCCCGCATTTTGGCTTACAATGGAC
 873▶ D  T  H  N  A  L  V  T  Q  F  K  P  D  S  K  N  Y  A  R  I  L  A  Y  N  G

2696 GCTTCAATGGCCCACCCAGTGCCGTCATCGACTTCGATACTCCGGAGGGTGTACCATCGCCGGTTCAGGGACTGGAT
 899▶R  F  N  G  P  P  S  A  V  I  D  F  D  T  P  E  G  V  P  S  P  V  Q  G  L  D

2773 GCCTATCCTCTGGGCTCCTCGGCCTTCATGCTCCACTGGAAGAAGCCGCTGTATCCCAATGGCAAGCTCACTGGCTA
 925▶A  Y  P  L  G  S  S  A  F  M  L  H  W  K  K  P  L  Y  P  N  G  K  L  T  G  Y

2850 CAAGATCTACTACGAGGAGGTTAAGGAGAGCTATGTGGGCGAGCGACGCGAATACGATCCACACATCACCGATCCCA
 950▶  K  I  Y  Y  E  E  V  K  E  S  Y  V  G  E  R  R  E  Y  D  P  H  I  T  D  P
```

```
2927 GGGTCACACGCATGAAGATGGCCGGCCTGAAGCCCAACTCCAAGTACCGCATCTCCATCACTGCCACCACGAAAATG
 976 R  V  T  R  M  K  M  A  G  L  K  P  N  S  K  Y  R  I  S  I  T  A  T  T  K  M

3004 GGCGAGGGATCTGAACACTATATCGAAAAGACCACGCTCAAGGATGCCGTCAATGTGGCCCCTGCCACGCCATCTTT
1002 G  E  G  S  E  H  Y  I  E  K  T  T  L  K  D  A  V  N  V  A  P  A  T  P  S  F

3081 CTCCTGGGAGCAACTGCCATCCGACAATGGACTAGCCAAGTTCCGCATCAACTGGCTGCCAAGTACCGAGGGTCATC
1027  S  W  E  Q  L  P  S  D  N  G  L  A  K  F  R  I  N  W  L  P  S  T  E  G  H

3158 CAGGCACTCACTTCTTTACGATGCACAGGATCAAGGGCGAAACCCAATGGATACGCGAGAATGAGGAAAAGAACTCC
1053 P  G  T  H  F  F  T  M  H  R  I  K  G  E  T  Q  W  I  R  E  N  E  E  K  N  S

3235 GATTACCAGGAGGTCGGTGGCTTAGATCCGGAGACCGCCTACGAGTTCCGCGTGGTGTCCGTGGATGGCCACTTTAA
1079 D  Y  Q  E  V  G  G  L  D  P  E  T  A  Y  E  F  R  V  V  S  V  D  G  H  F  N

3312 CACGGAGAGTGCCACGCAGGAGATCGACACGAACACCGTTGAGGGACCAATAATGGTGGCCAACGAGACGGTGGCCA
1104  T  E  S  A  T  Q  E  I  D  T  N  T  V  E  G  P  I  M  V  A  N  E  T  V  A

3389 ATGCCGGATGGTTCATTGGCATGATGCTGGCCCTGGCCTTCATCATCATCCTCTTCATCATCATCTGCATTATCCGA
1130 N  A  G  W  F  I  G  M  M  L  A  L  A  F  I  I  I  L  F  I  I  I  C  I  I  R

3466 CGCAATCGGGGCGGAAAGTACGATGTCCACGATCGGGAGCTGGCCAACGGCCGGCGGGATTATCCCGAAGAGGGCGG
1156 R  N  R  G  G  K  Y  D  V  H  D  R  E  L  A  N  G  R  R  D  Y  P  E  E  G  G

3543 ATTCCACGAGTACTCGCAACCGTTGGATAACAAGAGCGCTGGTCGCCAATCCGTGAGTTCAGCGAACAAACCGGGCG
1181  F  H  E  Y  S  Q  P  L  D  N  K  S  A  G  R  Q  S  V  S  S  A  N  K  P  G

3620 TGGAAAGCGATACTGATTCGATGGCCGAATACGGTGATGGCGATACAGGCATGAATGAAGATGGATCCTTTATTGGC
1207 V  E  S  D  T  D  S  M  A  E  Y  G  D  G  D  T  G  M  N  E  D  G  S  F  I  G

3697 CAATATGGACGCAAAGGACTTTGA
1233 Q  Y  G  R  K  G  L
```